CONSTRAINING GLOBAL BIOGENIC EMISSIONS AND EXPLORING SOURCE CONTRIBUTIONS TO TROPOSPHERIC OZONE: MODELING APPLICATIONS

A Dissertation
Presented to
The Academic Faculty

By

Changsub Shim

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in the
School of Earth and Atmospheric Sciences

Georgia Institute of Technology

August, 2006

Constraining Global Biogenic Emissions and Exploring Source Contributions to Tropospheric Ozone: Modeling Applications

Approved by:

Dr. Yuhang Wang, Advisor
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Dr. Athanasios Nenes
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Dr. Derek Cunnold
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Dr. Serge Guillas
School of Mathematics
Georgia Institute of Technology

Dr. Rodney Weber
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Date Approved: June 13, 2006

To my lovely family and their love and support

ACKNOWLEDGEMENTS

I deeply appreciate tremendous support from many persons during my Ph.D. experience. I thank my advisor, Dr. Yuhang Wang for his long-term support and guidance with his scientific enthusiasm. My group members, Tao Zeng, Yasuko Yoshida, Yunsoo Choi, Wei Liu, Rita Deng, and Chun Zhao, gave me encouragement and valuable suggestions. I thank Dr. Derek Cunnold, Dr. Rodney Weber, Dr. Athanasios Nenes, and Dr. Serge Guillas for taking the time to serve on my thesis committee and for helpful discussions along the way. My research has also been assisted by Dr. Kelly Chance, Dr. Paul Palmer, Dr. Alex Guenther, Dr. Ronald Blake, Dr. Mark Jacobson, Dr. Daniel Jacob, Bob Yantosca, Dorian Abbot, and many other persons. I also thank many former and current colleges and friends at EAS GIT, particularly, Dr. YS Yang, Dr. CH Song, David Kim, GR Jung, SI Lee, and SW Kim. I thank Dr. Keon Kim at Korea University for his leading me to scientific field. Besides my research experience, I have gratefully enjoyed friendship with KW Kim, TH Jung, SN Park, YD Cho, SW Ahn, and JY Mun. I thank the Reverend John Seo, the Reverend Joseph Choi, and Antonio Hwang for their spiritual encouragement. Above all, I have been always blessed by the tremendous love, belief, support of my family: my parents and my sister.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AI	aerosol index
AMF	air mass factor
ATSR	Along Track Scanning Radiometer
AVHRR	Advanced Very High Resolution Radiometer
^{7}Be	Beryllium - 7
BUV	Backscattered Ultraviolet (spectrometer)
CMB	chemical mass balance
CFC	chlorofluorocarbon
CMDL	Climate Monitoring and Diagnostics Laboratory (NOAA)
CTM	Chemical Transport Model
DOAS	Differential Optical Absorption Spectroscopy
ECMWF	European Centre for Medium-Range Weather Forecasts
ERS-2	European Remote Sensing-2 (satellite)
fCi/SCM	femtocurie per square centimeter ($1x10^{-15}$ Ci/SCM, an unit of ^{7}Be)
FREETEX	free tropospheric experiment
GEIA	Global Emissions Inventory Activity
GEOS-Chem	global Goddard Earth Observing System – Chemistry (model)
GMAO	Global Modeling Assimilation Office
GOME	Global Ozone Monitoring Experiment
GOMECAT	GOME cloud retrieval algorithm
HC	hydrocarbon
HCFC	hydrochlorofluorocarbon
HCHO	atmospheric formaldehyde
hPa	hecto Pascal
HY-SPLIT	hybrid single-particle Lagrangian intergrated trajectory
IIASA	international institute for applied systems analysis
INTEX-NA	intercontinental chemical transport experiment – North America
IPCC	Intergovernmental Panel on Climate Change
IR	infra red (spectrum range)
LAI	leaf area index
LIDORT	linear discrete ordinary radiative transfer
NMHC	nonmethane hydrocarbon
NOAA	National Oceanic and Atmospheric Administration
NOx	nitrogen oxides ($NO + NO_2$)
NOy	total reactive, oxidized nitrogen
OVOC	oxygenated volatile organic compounds
PAN	peroxyacetyl nitrate
PCA	principal component analysis
PEM-Tropics	Pacific Exploratory Mission in the Tropical Pacific
PMF	positive matrix factorization
ppbv	parts per billion by volume
ppmv	parts per million by volume

pptv	parts per trillion by volume
RTM	radiative transfer model
SAA	South Atlantic Anomaly
SCD	slant column density
SVD	singular value decomposition
SZA	solar zenith angle
TOA	top of the atmosphere
TOPSE	tropospheric ozone production about the spring Equinox
TRACE-P	TRAnsport of Chemical Evolution over the Pacific
3-D	three-dimensional
TOMS	Total Ozone Mapping Spectrometer
UV/VIS	ultra violet/visible (a range of spectrum)
VCD	vertical column density
VOC	volatile organic compound
θ	potential temperature

SUMMARY

Biogenic isoprene plays an important role in tropospheric chemistry. We use HCHO column measurements by the Global Ozone Monitoring Experiment (GOME) to constrain isoprene emissions. Using the global Goddrad Earth Observing System—Chemistry (GEOS-Chem) as the forward model, a Bayesian inversion of GOME HCHO observations from September 1996 to August 1997 is conducted. Column contributions to HCHO from 12 sources including 10 terrestrial ecosystem groups, biomass burning, and industry are considered and inverted for 8 geographical regions globally. The a posteriori solution reduces the model biases for all regions, and estimates the annual global isoprene emissions of 566 Tg C yr^{-1}, ~50% larger than the a priori estimate. Compared to the Global Emissions Inventory Activity (GEIA) inventory (~500 Tg C yr^{-1}), the a posteriori isoprene emissions are generally higher at mid latitudes and lower in the tropics. This increase of global isoprene emissions significantly affects tropospheric chemistry, decreasing the global mean OH concentration by 10.8% to 0.95×10^6 molecules/cm^3. The atmospheric lifetime of CH_3CCl_3 increases from 5.2 to 5.7 years.

Positive matrix factorization (PMF), an advanced method for source apportionment, is applied to TRAnsport of Chemical Evolution over the Pacific (TRACE-P) measurements and it is found that cyanogenesis in plants over Asia is likely an important emission process for CH_3COCH_3 and HCN. This approach also is applied to estimate source contributions to the tropospheric ozone (O_3) with Tropospheric Ozone Production about the Spring Equinox (TOPSE) and TRACE-P measurements. The corresponding GEOS-Chem simulations are applied to the same factor-projected space in order to evaluate the model simulations. Intercontinental transport of pollutants is most responsible for increasing trend of springtime O_3, while stratospheric influence is the largest contributions to troposperic O_3 variability at northern middle and high latitudes. On the other hand, the overall tropospheric contributions to O_3 variability are more important at northern low latitudes by long-range transport, biomass burning, and industry/urban emissions. In general, the simulated O_3 variabilities are comparable with those of observations. However, the model underestimates the trends of and the contributions to O_3 variability by long-range transport of O_3 and its precursors at northern middle and high latitudes.

CHAPTER I

INTRODUCTION

1.1 Context and Motivation

Global atmospheric ozone (O_3) constitutes a critical component of global atmosphere and environment. The thick stratospheric O_3 layer and its photolysis protect living organs from harmful ultra violet (UV) radiation from the sun. The majority of atmospheric O_3 exists in the stratosphere, the source of which can be explained by the Chapman Mechanism through the photolysis of atmospheric oxygen (O_2). Its vertical structure gives stratospheric stability. Stratospheric O_3 affects upper tropospheric O_3 significantly through dynamic exchange processes between the upper troposphere and the lower stratosphere (Holton, 1995). The role of O_3 is also important in troposphere because photolysis of O_3 with abundant water vapor (H_2O) in troposphere produces the hydroxyl radical (OH), which is the most important oxidant in troposphere. This tropospheric oxidation by OH is the dominant removal process of most atmospheric gases because the troposphere accounts for 85% of atmospheric mass, which determines the lifetime of greenhouse gases such as methane (CH_4) as well. Tropospheric O_3 is also a major greenhouse gas (IPCC, 2001).

Tropospheric O_3 near the surface, however, is harmful to living cells because of its strong oxidizing capacity. Exposure to high concentrations of O_3 may cause human respiratory disease such as asthma and cause damage to crop production (Koren, 1995; Fuhrer, 2002). The source of tropospheric O_3 includes photochemical production, which is affected by the emissions of and transport of O_3 precursors, and transport from the

stratosphere. The photochemical O_3 production in the troposphere is rather complicated. This production mechanism involves many precursors including nitrogen oxides (NO_x) and volatile organic compounds (VOCs). Figure 1.1 illustrates photochemical O_3 productions in the troposphere. Tropospheric O_3 concentration is controlled by many environmental factors such as time, geographical locations (which are affected by meteorological and topological conditions), and emissions of NO_x and VOCs, etc.

It is difficult to estimate the exact amount of global or regional OH and O_3 burden not only because of the complex relationships between these meteorological and topological factors, but also because of uncertainties in the emissions of O_3 precursors. In particular, there is limited quantitative knowledge about biogenic VOC emissions and their large-scale atmospheric impact.

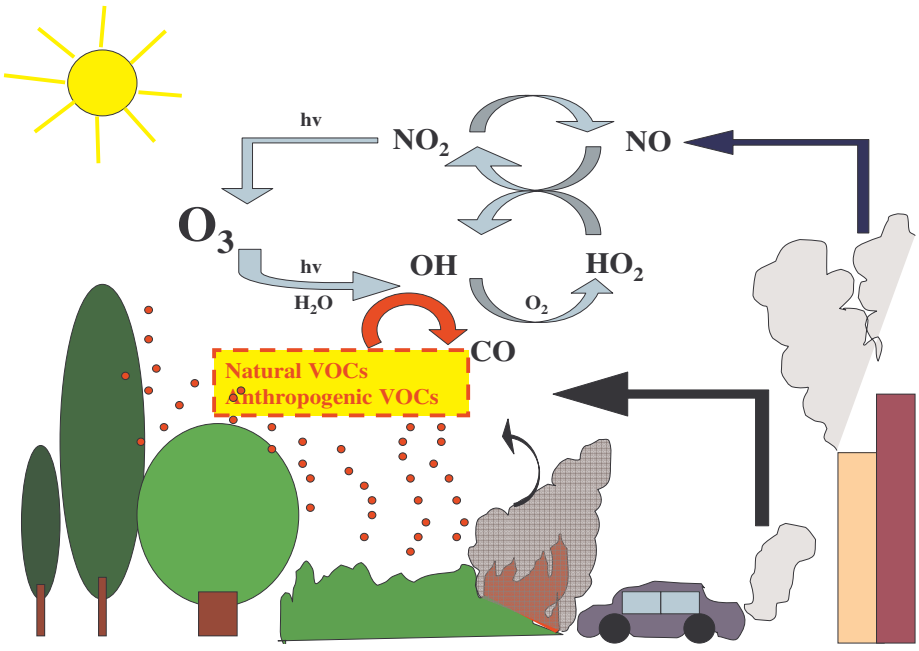


Figure1.1 Tropospheric photochemical ozone production mechanism controlled by various VOC and NO_x emissions during the day time.

In order to better understand tropospheric O_3 and photochemistry, this dissertation focuses on two major areas.

(1) Natural VOC emissions.

Natural VOCs, including isoprene, the dominant biogenic VOC, account for more than one third of global VOC emissions (Guenther *et al*., 1995). However, these natural VOC emissions (e.g., global isoprene) estimates are highly uncertain due to a lack of direct measurements. In addition to the effect on tropospheric O_3, the biogenic VOCs might contribute to the tropospheric aerosol budget significantly by formation of secondary organic aerosols (SOAs) (Claeys *et al*., 2004; Heald *et al*., 2005). Atmospheric VOCs not only affect regional air quality, but also are linked to global climate change driven by greenhouse gases and organic aerosols. A recent advancement is to constrian global isoprene emissions using the inverse method based on 3-D chemical transport model and satellite HCHO (formaldehyde) measurements. In addition, there is growing interest in major oxygenated VOCs (OVOCs), such as acetone ($CH_3C(O)CH_3$), methanol (CH_3OH), and acetaldehyde (CH_3CHO) because their abundance in the remote troposphere has been found by in-situ measurements (e.g., Sing *et al*. 2000, 2001). This dissertation uses factor analysis to conduct the source apportionments of the OVOC measurements and discusses the contributions of biogenic processes to acetone, acetaldehyde, and hydrogen cyanide.

(2) Sources of tropopshreic O_3 variability and its springtime increase in the northern hemisphere.

The sources of tropospheric O_3 involve either photochemical production coupled with transport within the troposphere or are transport from the stratosphere. The correct

knowledge of source contributions to tropospheric O_3 is of particular importance in global atmospheric chemistry and climate change. Photochemical and dynamic environments are different with geographical locations (e.g., latitudes). At mid to high latitudes, a significant increase of tropospheric O_3 is observed in the springtime, and many studies have tried to attribute this O_3 enhancement (e.g., Logan, 1985; Liu *et al*., 1987; Oltmans and Levy, 1992; Dibb *et al*., 1994; Wang *et al*., 1998, 2003a; Mauzerall *et al*., 2000).

Tropospheric Ozone Production about the Spring Equinox (TOPSE, February – May, 2000) aircraft measurements were conducted to investigate the springtime O_3 maximum at mid to high latitudes over North America. Here we specifically focused on the scientific questions on TOPSE and applied an advanced factor analysis to those measurements to understand the main contributors to O_3 variability and its springtime increase. With the same methods, the contributions of Asian outflow to O_3 variability at mid to low latitudes were analyzed with measurements during TRAnsport of Chemical Evolution over the Pacific (TRACE-P, March – April, 2001).

1.2. Methodology

1.2.1. Remote Sensing and Retrieval of Atmospheric HCHO

Atmospheric constituents interact with electromagnetic waves. The solar radiation is the most important life-supporting electromagnetic waves on earth with an energy flux of 1367 Wm^{-2} (at the top of the atmosphere (TOA)). The interactions between electromagnetic waves and atmospheric molecules include absorption (radiation converts into the internal energy of molecules), transmission (radiation transmits without

interactions), reflection (radiation changes direction), and scattering (radiation changes its phase or direction with/without change of energy). The absorbed radiation into constituents increases their internal energy in forms of translational (unquantized motion to increase kinetic energy), rotational (quantized with the smallest energy change), vibrational (quantized spectrum covers mostly infrared/near-infrared wavelength (1— 20 μm)), and electronic (quantized with the greatest energy change, causing excitation of the outer-shell electron(s)) energies.

The atmospheric constituents absorb at discrete frequencies corresponding to their transitions of energy levels. Vibrational and rotational transitions are of primary interest for absorption and emissions of terrestrial radiation in the atmosphere. Many of the major constituents such as H_2O, CH_4, CO_2, and O_3 have vibrational and rotational bands mostly within the thermal infrared range of the spectrum (Figure 1.2). Every constituent has its unique absorption bands due to its unique energy transitions by its absorption cross section (the ability of a particular molecule to absorb a photon at a particular wavelength). That uniqueness is basically applied to retrieve concentrations of remotely-measured atmospheric constituents.

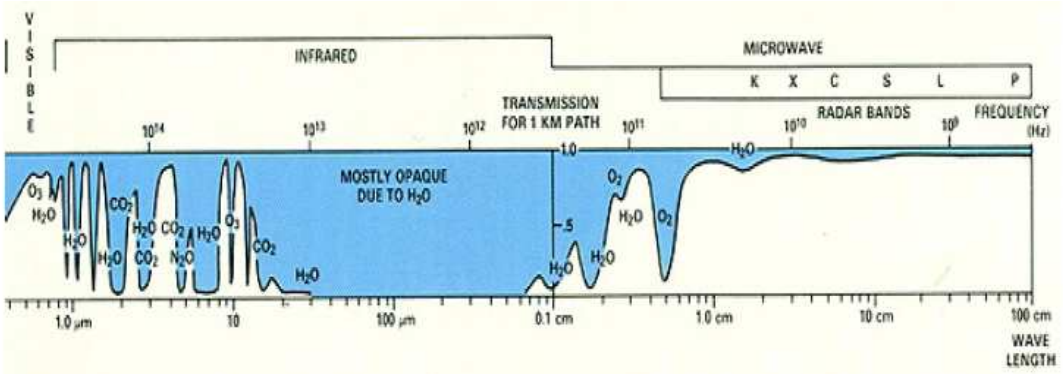


Figure 1.2. **A diagram showing relative atmospheric radiation transmission/ absorption at different wavelengths. Blue zones show low passage of incoming and/or outgoing radiation due to the absorptions of H_2O and other constituents, and white areas denote atmospheric windows where the radiation does not interact much with air molecules (Sokolik, 2004).**

Atmospheric formaldehyde (HCHO) is the tracer of interest in this dissertation (Chapter 2). The wavelength of HCHO absorption is within the ultra-violet region (UV-region, 335 – 355 nm, Figure 1.3). A satellite instrument measures UV backscattering spectra. In this dissertation, the satellite HCHO measurements from the Global Ozone Monitoring Experiment (GOME) are used to constrain global isoprene emissions (Chapter 2). GOME is on board the ERS-2 satellite launched in 1995 to monitor atmospheric trace gases. It measures solar backscattered radiances in nadir-viewing mode, with spectral resolutions of 0.2 – 0.4 nm within UV – near IR range (Burrows *et al*., 1999). The satellite moves in a descending node at a sun-synchronous orbit, passing the equator at 10:30 a.m. (local time). The information that can be retrieved from the UV backscattering spectra includes reflected solar spectrum, surface/cloud albedo, atmospheric scattering, temperature, pressure, and concentration of absorbing chemical tracers.

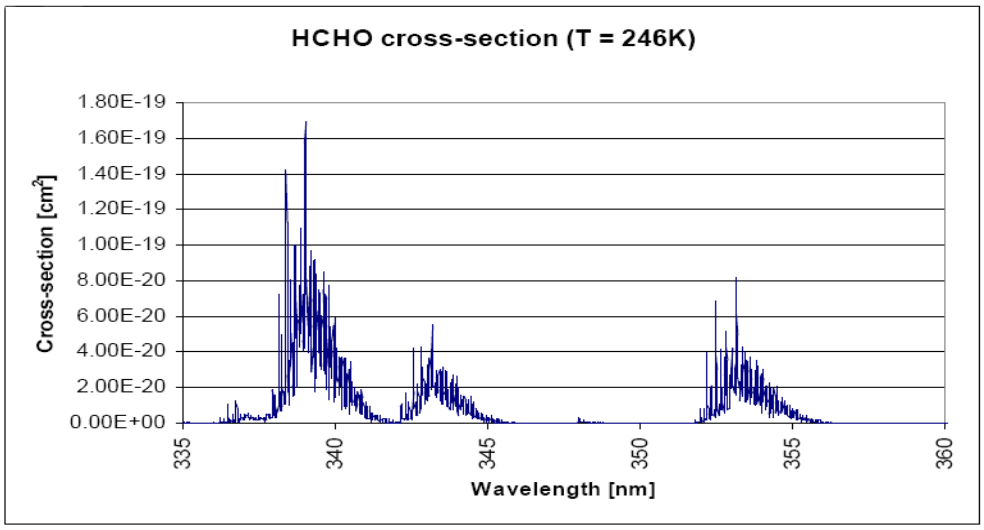


Figure 1.3. Absorption cross section of atmospheric HCHO (Cantrell *et al*., 1990).

Obtaining SCD

After calibration of the detected spectra, slant column density (SCD) can be derived using the Beer-Lambert Law if one knows the values of absorption cross sections. In practice, the slant column density (SCD) of tracer is then determined by least squares fitting a linear combination of reference absorption cross-section spectra of trace gases and a Ring reference spectrum to the measured optical density in the UV/ VIS (ultra violet / visible) (e.g., Platt, 1994; Chance *et al*., 2000).

Conversion from SCD to VCD

In order to obtain a vertical column density (VCD), the dependence of SCD in terms of the viewing angle with latitudes should be considered. Air mass factor (AMF) is used to convert SCD to VCD. AMF depends on the vertical distribution of chemical species because of atmospheric scattering. In particular, HCHO is mostly distributed within the boundary layer, where the scattering is most important. Palmer *et al*. (2001) formulates

AMF as the integral of the relative vertical distribution factor (shape factor) of the chemical species over the depth of the atmosphere, weighted by an altitude-dependent degree of scattering (scattering weights) using a radiative transfer model (RTM) and a chemical transport model (CTM). HCHO vertical columns in Chapter 2 are retrieved based on those procedures.

1.2.2. Global Chemical Transport Model.

Goddard Earth Observing System–Chemistry (GEOS-Chem) is used as the global CTM for this dissertation. Tropospheric chemistry is driven by assimilated meteorological data from the Global Modeling Assimilation Office (GMAO) (Schubert *et al*., 1993). The 3-D meteorological data are updated every 6 hours; mixing depths and surface fields are updated every 3 hours. For meteorological fields, we used GEOS-STRAT (4° x 5° horizontal and 26 vertical resolution, Chapter 2), GEOS-3 (2° x 2.5° horizontal and 30 vertical resolution, Chapter 5) for this dissertation. The vertical levels extend up to approximately 0.1 hPa. A comprehensive description and evaluation of the GEOS-Chem model for tropospheric O_3–NO_x–VOC chemistry is given by Bey *et al*. (2001).

The chemical mechanism is that of Horowitz *et al*. (1998) with updates of more recent kinetics data. It includes a detailed representation of oxidation pathways for five non-methane hydrocarbons (NMHCs) (ethane, propane, lumped >C3 alkanes, lumped >C2 alkenes, and isoprene). The organic peroxides produced from isoprene oxidation are recycled by photolysis and reaction with OH (Lurmann *et al*., 1986). Numerical integration of the mechanism is done with a fast Gear solver (Jacobson and Turco, 1994). Photolysis rates are computed using the Fast-J radiative transfer algorithm (Wild *et al*.,

2000), which includes Rayleigh scattering as well as Mie scattering by clouds. Monthly averaged UV albedo fields are used from the work by Herman and Celarier (1997). Aerosol scattering is included in the model by specifying aerosol optical depths (AODs) from a correlation with surface O_3, as described by Fiore *et al*. (2002).

Gridded emission fields for anthropogenic NO_x and NMHCs in 1985 are taken from the study of Wang *et al*. (1998) and are scaled to later years using updated measurements or emission estimates. Biogenic hydrocarbon emissions include isoprene as discussed in the Chapter 2, and additional contributions from acetone and propene. Emissions of terpenes and methanol are not included; as will be discussed in Chapter 2, they are unlikely to make significant contributions to the HCHO column signal from GOME. The details of current version of GEOS-Chem are available at http://www-as.harvard.edu/chemistry/trop/geos/doc/man/index.html.

1.2.3. Inverse Modeling

Inverse modeling in atmospheric tracers involves making statistical inferences of physical parameters ("causes," e.g., tracer sinks or sources) derived from the measurements ("effects," e.g., tracer concentrations) (Enting, 2002). In contrast to inverse modeling, forward modeling estimates tracer concentrations based on current knowledge of the parameters (a priori parameters). Figure 1.4 illustrates the concept of inverse modeling. The inverse modeling in this dissertation focuses on deducing better isoprene emissions (a posteriori emissions) from GOME HCHO observations. GEOS-Chem chemical transport model is employed as the forward model driven by a priori parameters regarding isoprene emissions, contributing to HCHO concentrations during the growing season. We used Bayesian least squares methods to obtain a posteriori global isoprene

emissions, which better fit the GOME HCHO columns. The detailed descriptions of the inverse modeling are in Chapter 2.

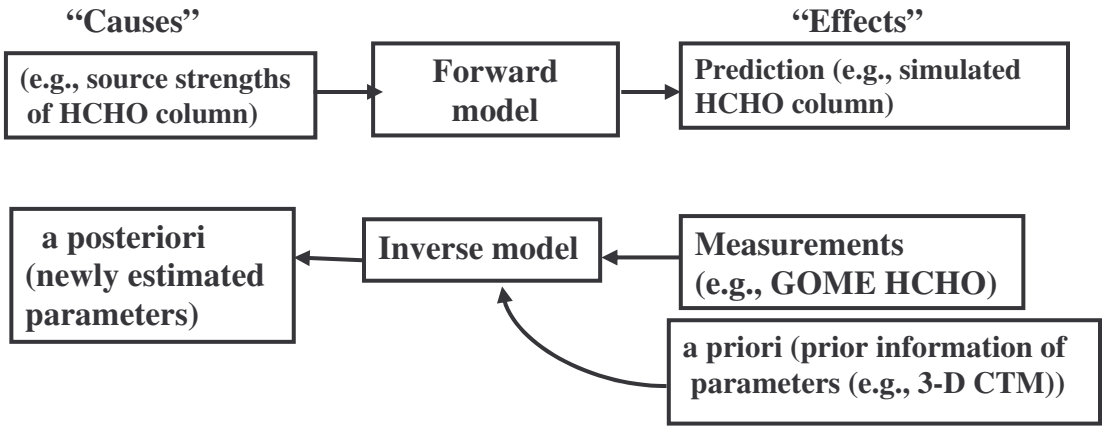


Figure 1.4. Schematic diagrams illustrating forward model and inverse model.

1.2.4. Receptor Modeling

Analyzing air quality requires various types of knowledge: sources types and emission rates, transport of constituents, and chemical and physical transformations of the constituents (e.g., photochemistry). The mathematical source-oriented model (e.g., 3-D CTM) uses all of these types of knowledge to predict air quality. For correct predictions, a survey of emissions over the studied regions and reasonable meteorological data are necessary. Field measurements are often used to evaluate the model results. The major limitations in the source-oriented model are associated with the variety of uncertainties from those parameterizations. Therefore, it is useful to have an alternative approach to identify sources and their contributions to ambient trace gases. One such approach is called receptor (-oriented) modeling (Wei, 2002).

Receptor models are based on measured mass concentrations and the use of appropriate mass balances. The chemical mass balance assumes that the total concentration of the chemical tracer A is the sum of contributions from a number of independent sources.

$$A_{total} = A_{source1} + A_{source2} + A_{source3} + \ldots \quad (1.2.4.1)$$

Every independent source emits not only substance A, but also other constituents such as B, C, etc. Thereby, $A_{source1}$, for example, can be expressed as

$$A_{source1} = g_{source1} f_{A,source1} \quad (1.2.4.2)$$

where $A_{source1}$ is the atmospheric concentration of A from source1, $g_{source1}$, which is the mass concentration of source1 in the atmosphere, and $f_{A,source1}$ is the gravimetric concentration or mass fraction of A in source1. Considering other sources (p independent sources), one can rewrite the equation as

$$A_{total} = g_{source1} f_{A,source1} + g_{source2} f_{A,source2} + \ldots + g_{sourcep} f_{A,sourcep} \quad (1.2.4.3)$$

More generally, the mass balance formula above can be rewritten accounting for m chemical species in n samples:

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} \quad (1.2.4.4)$$

where x_{ij} is the concentration of the j^{th} species in the i^{th} sample, g_{ik} is the mass concentration constituents from the k^{th} source contributing to the i^{th} sample, f_{kj} is the gravimetric concentration (fraction) of the j^{th} species in constituents emitted from the k^{th} source.

There are various methods one might use to solve the equation 1.2.4.4. First, the Chemical Mass Balance (CMB) Model fits each of the known source profiles (f_{kj}) toward the measured data x_{ij}, resulting in the source contribution g_{ik} (Hopke, 1985). The problem of CMB is that the source profile is not always well known, which may lead to incorrect source estimates. Second, multivariate receptor models are based on the idea that the time dependence of a chemical species at the receptor site will be the same for species from the same source. Species of similar variability are grouped together in a minimum number of factors (sources) that explain the variability of the data set, which is often called factor analysis. This analysis assumes that highly correlated compounds come from the same source. The common example of multivariate analysis is Principal Component Analysis (PCA). The results of that methods are, however, often ambiguous and difficult to interpret in view of physical sciences, since the resolved profiles of a source may include negative variabilities.

1.2.4.1. The PMF Method

Pattero and Tapper (1993, 1994) provided an advanced factor analysis obtaining a better identification of source contributions to observed particle compositions. This method basically uses the error estimate of individual measurements to provide optimum data scaling and permits a better treatment of missing and below detection-limit values. More importantly, requiring non-negativity for both factor components (to be explained

in the following section) renders the solutions easier to interpret and allows the resulting features to make physical sense. This approach is called positive matrix factorization (PMF). PMF has been successfully applied to the analysis of many kinds of researches (e.g., Polissar *et al*., 1998; Liu *et al*., 2003; 2005; Wang *et al*., 2003a; Xie *et al*., 1999). The PMF method is applied as an important analysis tool for source apportionments in this dissertation (Chapters 3, 4, and 5).

Let's suppose X is a (n x p) data matrix consisting of the measurements of p chemical species in n sample observations. The objective of PMF modeling is to determine the number of sources, the chemical composition profile of each source, and the amount that each source contributes to the sample. The corresponding two-way factor analysis model can be written as follows.

$$\mathbf{X = GF + E} \qquad (1.2.4.5)$$

where G is n x k matrix and F is k x p matrix.

or

$$x_{ij} = \sum_{k=1}^{l} g_{ik} f_{kj} + e_{ij} \qquad (1.2.4.6)$$

where i = 1,…, n ; j=1,…,p; k=1,…,l. G is a (n x k) matrix of source contributions (factor scores) to the samples and F is a (k x p) matrix of source chemical compositions (factor loadings or profile). Each sample is an observation along the time axis, so G describes the temporal variation of the sources. E represents the part of the data not explained by PMF, which comes from residuals or errors (Paatero and Tapper, 1994).

In the PMF model, the solution is a weighted least squares fit, where the known standard deviations for each value of "x" are used for determining the weights of the residuals in matrix "E." The data for maximum likelihood estimates are assumed to be normally distributed. PMF constrains factor loadings and factor scores to nonnegative values and thereby minimizes the ambiguity caused by rotating factors. This is one of the major differences between PMF and principal component analysis (PCA).

The objective of PMF is to minimize the sum of the weighted residuals (Q) defined as

$$Q = \sum_{i=1}^{n}\sum_{j=1}^{p}\frac{(x_{ij} - g_{ik} f_{k\,i})^2}{\sigma_{ij}^{\,2}} = \sum_{i=1}^{n}\sum_{j=1}^{p}\left(\frac{e_{ij}}{\sigma_{ij}}\right)^2 \qquad (1.2.4.7)$$

where $\sigma_{i,j}$ is a user-defined error estimate for each elements of x_{ij} in a matrix. Theoretically, if there are no outliers, then the final Q value after iterations must be approximately same as the number of data points in the original matrix X. The theoretical Q value is based on the "degrees of freedom" and equals the number of data points in the array. However, the actual calculated Q value is frequently much bigger than the theoretical one. Such a large difference between theoretical and actual Q values can be an indication that the data detection limits are underestimated (Paatero, 2000).

In PMF, non-negativity constraints are imposed on both the source profiles (F) and the source contributions (G). All of the species concentrations must be greater than 0, and each sample must have each source contribution greater than 0. The PMF algorithm then calculates G and F matrices that minimize the value of the variable Q with $g_{ik} \geq 0$ and $f_{kj} \geq 0$ for k = 1, .. , p. The problem to obtain optimized Q values cannot be solved by

singular value decomposition (SVD). It can be solved, however, by iterative algorithms for optimization. This algorithm minimizes the Q function with respect to the unknown matrix, presetting either the G or F matrix as known. Then the recently solved matrix is taken as known, and the other is solved until it reaches convergence. The details of the algorithm are described in Paattero and Tapper (1994). Thus, this is a least-squares problem with the values of G and F to be determined.

An important parameter in PMF analysis in this dissertation is the explained variation (EV), which is a magnitude of the contribution of each chemical species for each identified factor (Lee *et al*., 1999). EV values can be considered the relative variation of elements at F matrix. The EV value of chemical species j in the k^{th} factor can be calculated by

$$EV_{kj} = \frac{\sum_{i=1}^{n} \frac{\left| G_{ik} F_{kj} \right|}{\sigma_{ij}}}{\sum_{i=1}^{n} \frac{\left(\sum_{k=1}^{l} \left| G_{ik} F_{kj} \right| + \left| e_{ij} \right| \right)}{\sigma_{ij}}} \qquad (1.2.4.8)$$

Cumulative variance of EV explains how much of each species is explained in PMF analysis. For instance, if cumulative variance of O_3 is 0.8, it can be said that approximately 80% of total O_3 variation is explained by identified factors.

Choosing an appropriate number of factors

During PMF analysis, it is important to choose the number of factors that provide physically meaningful results. Basically, PMF computes the factors in random order if random values are used for starting the computation (Paatero, 2000). In this analysis, the

order factor is determined by sorting the center-of-mass locations of the G or F matrix in ascending order. A fundamental condition for choosing the number of factors is that the maximum number of factors should be less than the number of tracers. Assigning a smaller number of factors than actual amount of factors in a matrix causes a factor that represents mixed-source characteristics. On the other hand, too many assigned factors cause residual factor(s) that represent(s) noise due to unexplained residuals.

There are mathematical ways to validate the number of factors. Lee *et al*. (1999) used the information from the scaled residual and rotational matrix to determine a reasonable range of factors. Monitoring the Q value with a number of factors also could inform maximum factor numbers. Despite the multiple possibilities of number of factors, the "best solution" must have the factors representing the most physical sense. Chapters 3, 4, and 5 explain more about choosing the number of factors.

Assumptions and limitations of PMF analysis

In spite of advanced optimizing algorithms and better physical constraints with non-negative values, there are limitations of receptor modeling when we try to estimate contributions to chemical tracers. First, the results of the receptor model are often sensitive to the selected suit of tracers, number of factors, assigned uncertainties, and matrix rotations, which could provide too many possible solutions. Second, since the source profiles are presumed to be added linearly, it is difficult to quantify the emissions for chemicals with large secondary productions. Third, the receptor model assumes that compositions of emissions basically do not change during transport from the point of emission to the point of measurement. That assumption also makes it difficult to measure the correct contributions of emissions because transport processes can mix multiple

emissions, which causes reduced or mixed source characteristics. Those problems are addressed in Chapter 3. However, these last two limitations are less relevant to the source of O_3 (as discussed in Chapters 4 and 5) because there is no primary emission source for O_3; all the O_3 sources are either from secondary productions or transport. Thus the limitations do not directly affect the objectives of these last two chapters, which involve identifying the source characteristics and contributions, not estimating the primary emissions of chemicals. Despite its problems and limitations, the PMF analysis is still advantageous over source-oriented model studies when the source types, strength, and distributions are not well understood.

.

1.3 Scope of this Dissertation.

This dissertation analyzes the diverse source contributions to tropospheric O_3, particularly linked to terrestrial biogenic VOC emissions. Specifically, the chapters are organized as follows.

- **Chapter 2, "Constraining global isoprene emissions with Global Ozone Monitoring Experiment (GOME) formaldehyde column measurements."**

  The global isoprene emissions are constrained with GOME HCHO columns. HCHO is a good proxy for terrestrial isoprene emissions during the growing season. 3-D CTM (GEOS-Chem) is applied as a forward modeling. Twelve different pre-described contributions to isoprene emissions from 10 terrestrial ecosystems, biomass burning, and industry are applied to Bayesian inverse modeling. Constrained global isoprene emissions with GOME HCHO columns

are compared with Global Emissions Inventory Activity (GEIA) inventory and the resulting uncertainties of the emissions are greatly reduced.

- **Chapter 3, "Source characteristics of oxygenated volatile organic compounds and hydrogen cyanide."**

  Aircraft measurements (TRACE-P, PEM-Tropics B, and Intercontinental Chemical Transport Experiment – North America (INTEX-NA)) are analyzed to identify the source characteristics of OVOCs and hydrogen cyanides using the PMF method. The source characteristics are investigated and their contributions are discussed with implications for their global source budgets.

- **Chapter 4, "Intercontinental transport of pollution manifested in the variability and seasonal trend of springtime O_3 at northern middle and high latitudes."**

  The tracer measurements during TOPSE are analyzed with the PMF method in order to identify and estimate the contributions to springtime O_3 variability at northern middle and high latitudes. O_3 contributions from photochemistry and transport within the troposphere and from the stratosphere are of particular interest. This chapter presents the first PMF analysis to estimate those contributions.

- **Chapter 5. "Evaluation of model simulated source contributions to tropospheric ozone with aircraft observations on the factor-projected space."**

The chemical transport model (CTM) is usually evaluated by comparisons of the simulated tracer's concentrations with those of measurements. However, that type of evaluations cannot be used to examine how the model simulates the contributions of various processes to O_3. We apply the PMF method to examine the observed and simulated covariance structures of O_3 and its precursors. We use the CTM (GEOS-Chem) simulations and aircraft measurements (TOPSE and TRACE-P) for this purpose. The analysis results indicate that it is a viable method for evaluating model's capability in simulating process-based contributions to tropospheric O_3.

CHAPTER II

CONSTRAINING GLOBAL ISOPRENE EMISSIONS WITH GOME FORMALDEHYDE COLUMN MEASUREMENTS[1]

2.1 Introduction

Volatile organic compounds (VOCs) play an important role in oxidation chemistry in the troposphere (Chameides *et al*., 1992; Moxim *et al*., 1996; Houweling *et al*., 1998; Wang *et al*., 1998b; Poisson *et al*., 2000). Biogenic emissions are major sources of VOCs (e.g., Zimmerman, 1979; Lamb *et al*., 1987; Mueller, 1992). Isoprene (C_5H_8), in particular, represents almost half of the total source of biogenic VOCs, and almost 40% of total VOC emissions on a global scale (Guenther *et al*., 1995). Furthermore, the formation of secondary organic aerosols via photooxidation of isoprene affects the global climate (Limbeck *et al*., 2003; Claeys *et al*., 2004).

Isoprene emissions depend on vegetation types, light intensity, temperature and leaf area index (LAI) (Lamb *et al*., 1987; Guenther *et al*., 1995). Global emissions of isoprene are generally estimated by extrapolating from limited laboratory and field measurements to the prescribed global ecosystems. Various emission parameterizations have been proposed (e.g., Lamb *et al*., 1987; Guenther *et al*., 1995). Despite these efforts, large uncertainties still remain in the estimates (Hewitt and Street, 1992; Guenther *et al*., 1995).

---

[1]This chapter is extension of "Constraining global isoprene emissions with GOME formaldehyde column measurements," published at *Journal of Geophysical Research* in December 2005 (*110*, D24301, doi:10.1029/2004JD00529). Authors are **Changsub Shim**, Yuhang Wang, Yunsoo Choi, Paul Palmer, and Kelly Chance.

The difficulty lies in the scarcity of direct measurements. The problem is most acute for tropical ecosystems (Guenther *et al*., 1995; Pierce *et al*., 1998), which collectively account for more than half of global isoprene emissions (Guenther *et al*., 1995).

Formaldehyde (HCHO) is a product of VOC oxidation. The main sinks of HCHO are photolysis and the reaction with atmospheric hydroxyl radical (OH), and its lifetime against oxidation (order of hours) is short enough not to be significantly affected by transport. Methane (CH_4) is an important HCHO source, but it is well mixed in the troposphere due to its long lifetime. Methane oxidation provides the background HCHO levels.

Previously, isoprene emissions over North America in summer have been derived using HCHO column measurements from the Global Ozone Monitoring Experiment (GOME) (Chance *et al*., 2000; Palmer *et al*., 2003a; Abbot *et al*., 2003). It was found that isoprene is the dominant contributor to HCHO over North America in the growing season; the enhancements above the CH_4-oxidation induced background levels are generally linear with local isoprene emissions over North America (Chance *et al*., 2000; Palmer *et al*., 2003a). Using that information, the seasonal and interannual variations in GOME HCHO columns over North America has been investigated (Abbot *et al*., 2003). Here we extend these previous studies to the global scale and we also explicitly consider isoprene emissions from 10 vegetation groups to capture the large difference in base emissions for different types of vegetations. The sources from biomass burning and industry are also treated separately.

In this work, we apply GOME observations of HCHO column from September 1996 to August 1997 to constrain global isoprene emissions. In order to obtain best estimations

of isoprene emissions, we use statistical inferences to fit the model simulated HCHO column concentrations toward GOME-observed HCHO column concentrations (inverse modeling). To minimize the effects of GOME measurement uncertainties, we selected 8 regions with high signal-to-noise ratios (the ratio of slant column to signal fitting error > 4). These regions are located over North America, Europe, East Asia, India, Southeast Asia, South America, Africa, and Australia.

Model parameters (state vector) considered for the HCHO sources include the oxidation of isoprene from 9 major vegetation groups; the 10^{th} group includes isoprene from all other vegetation types and biogenic VOCs other than isoprene. The other two HCHO sources considered are biomass burning (combined with biofuel burning) and industry. The sources include primary emissions of HCHO and secondary chemical production during the oxidation of other VOCs.

Uncertainties of model source parameters and GOME measurements are taken into account through Bayesian inverse modeling (Rodgers, 2000) to produce the a posteriori global isoprene emissions. The global GEOS-Chem chemical transport model (Bey *et al*., 2001) is used for the a priori estimate. We conduct for each region an inversion of 12 different source types using monthly mean observations during growing seasons. The effects of a posteriori change of isoprene emissions on global O_3 and OH are estimated.

2.2 HCHO as a Proxy for Isoprene Emissions.

2.2.1. General Oxidation Mechanism of Atmospheric VOCs

During the day time, the oxidation of a hydrocarbon (hereafter we use RH for hydrocarbon, R for an organic group of RH) is mostly initiated by OH.

$$RH + OH + O_2 \rightarrow RO_2 + H_2O \quad (2.1)$$

The RO_2 (organic peroxy radical) produced above reacts with NO (nitrogen oxide) to produce NO_2 (nitrogen dioxide) and an organic oxy radical (RO), or it may react with HO_2 (hydroperoxy radical) to produce organic hydroperoxide (ROOH).

$$RO_2 + NO \rightarrow RO + NO_2 \quad (2.2.1)\text{: high } NO_X \text{ condition.}$$

$$RO_2 + HO_2 \rightarrow ROOH + O_2 \quad (2.2.2)\text{: low } NO_X \text{ condition.}$$

NO_2 is photolyzed to produce O_3.

$$NO_2 + h\nu + O_2 \rightarrow NO + O_3 \quad (2.3)$$

ROOH from (2.2.2) may either reacts with OH or photolyzes and ultimately produce R'CHO (aldehyde; R' denotes reduced carbon from R), OH and H_2O. RO radical from (2.2.1) can react with atmospheric oxygen (O_2), thermally decompose, or isomerizes. The subsequent mechanisms are complicated, but generally carbonyl compounds and a HO_2 radical are produced. These are simply described as follows (Jacob, *et al*, 1999).

$$RO + O_2 \rightarrow R'CHO + HO_2 \quad (2.4)$$

$$HO_2 + NO \rightarrow OH + NO_2 \quad (2.5)$$

The net reaction so far is

$$RH + 4O_2 \rightarrow R'CHO + 2O_3 + H_2O \quad (2.6)$$

If R' group is H (hydrogen atom), then above reaction (2.6) produces HCHO (formaldehyde). Thus HCHO is a major intermediate during the oxidation of VOCs.

Aldehydes (R'CHO) react with OH or photolyze.

$$R'CHO + OH \rightarrow R'CO + H_2O \quad (2.7.1)$$

$$R'CHO + h\nu + O_2 \rightarrow R'CO + HO_2 \quad (2.7.2)$$

$$R'CHO + h\nu \rightarrow R'O + H_2 \quad (2.7.3)$$

The R'CO (organic acyl radical) reacts rapidly with O_2 to produce R'O and HO_2

$$R'CO + O_2 \rightarrow R'O + HO_2 \quad (2.8)$$

The decomposed organic group (R') reacts through from (2.4) to (2.8) and ultimately produces CO and HO_2.

$$CHO + O_2 \rightarrow CO + HO_2 \quad (2.9)$$

CO (carbon monoxide) is finally oxidized to relatively stable products, CO_2 and O_3.

$$CO + OH + O_2 \rightarrow CO_2 + HO_2 \quad (2.10)$$

$$HO_2 + NO \rightarrow OH + NO_2 \quad (2.11)$$

$$NO_2 + h\nu + O_2 \rightarrow NO + O_3 \quad (2.12)$$

The systematic diagram of above mechanism is shown in Figure 2.1. On the basis of the hydrocarbon oxidation mechanism, it is clear that VOCs play an important role on tropospheric O_3 and HO_X chemistry. Biogenic isoprene (CH_2=CH-C(CH_3)=CH_2, simply C_5H_8) has much more complicated reaction pathways, involving more than hundred reactants, but the outlines of the reaction mechanisms are similar to above reactions.

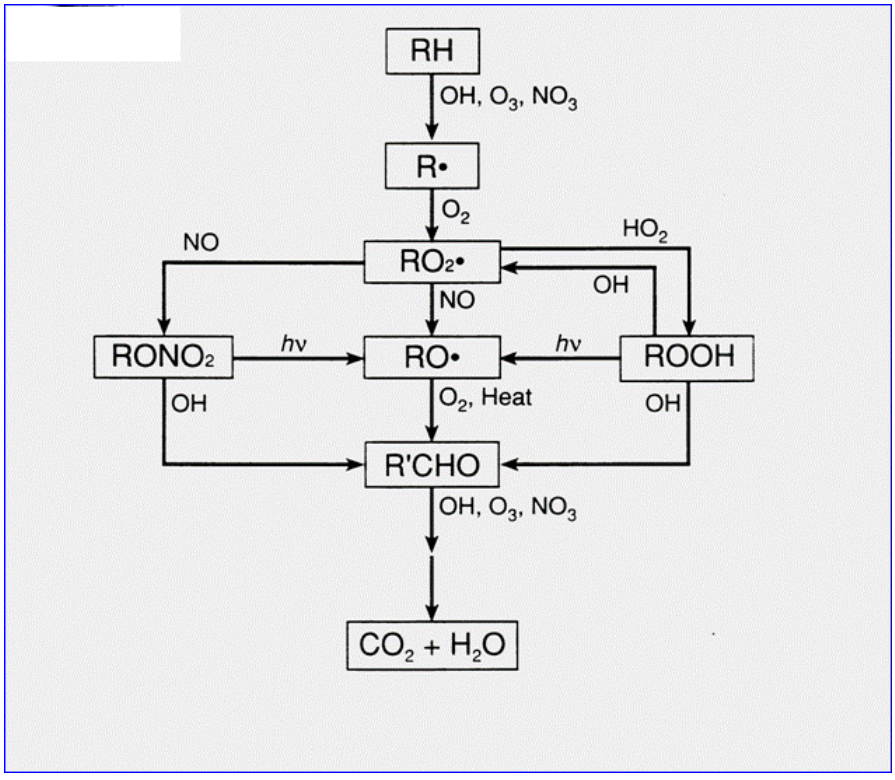


Figure 2.1. General oxidation pathway of atmospheric VOCs (Brasseur *et al.*, 1999).

2.2.2 HCHO as a Proxy for Isoprene Emissions.

The HCHO yield from isoprene oxidation on a per carbon basis is in the range of 0.3 – 0.45; it increases with NO_X concentrations (Horowitz *et al*., 1998; Palmer *et al*., 2003a). Palmer *et al*. (2003a) discussed the robustness of the isoprene-oxidation chemical mechanism in the model over North America in July 1996. They did not include the kinetics uncertainty in their inversion calculation. We assume in this work that this uncertainty in the estimated HCHO yield is small compared to GOME retrieval errors, which are fairly large (section 2.3). Quantitative assessment of the kinetics uncertainty with critical laboratory measurements is beyond the scope of this work.

Formaldehyde can be produced within an hour from isoprene emissions because the lifetime of isoprene is about 0.5 hour during late morning. The corresponding lifetimes of major secondary products of isoprene oxidation are about 1.5 – 2.5 hours (Carslaw *et al*., 2000). At GOME measurement time of 10:30 a.m. local time (LT), the impact of secondary products is also mitigated by the relatively weak isoprene emissions before the measurement time due to less light intensity and lower temperature. At surface wind speed of 0 – 10 m/s, the transport distance of secondary products is < 150 km since 6:30 am LT, much less than the model grid size (4°×5°). Thus, the relatively short lifetimes of isoprene, its major secondary products, and HCHO render the effect of transport insignificant in a coarse-resolution model. The isoprene oxidation with O_3 is insignificant because the reaction is much slower than that with OH and the HCHO yield is small (< 0.2) (Atkinson, 1994).

The per-carbon HCHO yields from larger biogenic VOCs, such as monoterpenes, are known to be much less than that of isoprene because of the efficient aerosol uptake of the

oxidation products (Kamens *et al*., 1982; Hatakeyama *et al*., 1991; Orlando *et al*., 2000; Palmer *et al*., 2003a). Methanol (CH_3OH), the other main biogenic HCHO source, has a much longer lifetime (several days). Formaldehyde from CH_3OH oxidation is distributed over large regions relative to the model grid size (Palmer *et al*., 2003a).

Industrial VOCs including alkanes, alkenes, and aromatics contribute less to HCHO than isoprene during growing seasons. The lifetimes of alkenes are generally longer than that of isoprene (Atkinson, 1994) and those of alkanes are much longer (Atkinson, 1994). Therefore HCHO from these VOCs are distributed over large regions. During growing seasons, their contributions to HCHO are relatively small over the regions with substantial biogenic isoprene emissions (to be shown in Table 2.3). The latter regions are the focus of our inverse modeling. The HCHO yields of aromatics are generally very small (less than a few percent) (Dumdei *et al*., 1988). Therefore the impact of those species is not important in this study.

Methane oxidation has a yield of about 1 HCHO per unit carbon and CH_4 is well mixed in the atmosphere because of its long lifetime (~10 years). Its contribution of about 30% to HCHO column defines the atmospheric HCHO background concentrations (Palmer et. al., 2003a). Recent studies (Palmer et. al. 2003a; Abbot *et al*., 2003) show that the variability of HCHO columns at northern mid latitudes reflects the production from isoprene oxidation. On a global scale, the variability of HCHO columns is also mostly affected by emissions and photochemistry over the regions of interest.

2.3 GOME HCHO Column Measurements and the Uncertainties

GOME is on board the ERS-2 satellite launched in 1995 to monitor atmospheric trace gases. It measures solar backscattered radiances with spectral resolutions of 0.2 – 0.4 nm in a relatively wide spectral range (237 – 794 nm, UV – near IR) (Burrows *et al*., 1999). The satellite moves in descending node passing the equator at 10:30 a.m. (local time) in a sun-synchronous orbit. In the nadir-viewing mode, it has a 32° across-track scan angle. It has a spatial resolution of 40×320 km^2 and takes ~3 days to cover the globe. Figure 2.2 illustrates the GOME scanning procedures. The HCHO absorption spectra at 337 – 356 nm are fitted to reference spectra by non linear least square method to determine slant columns with a fitting uncertainty of 4.0×10^{15} molecules cm^{-2} (Chance *et al*., 2000).

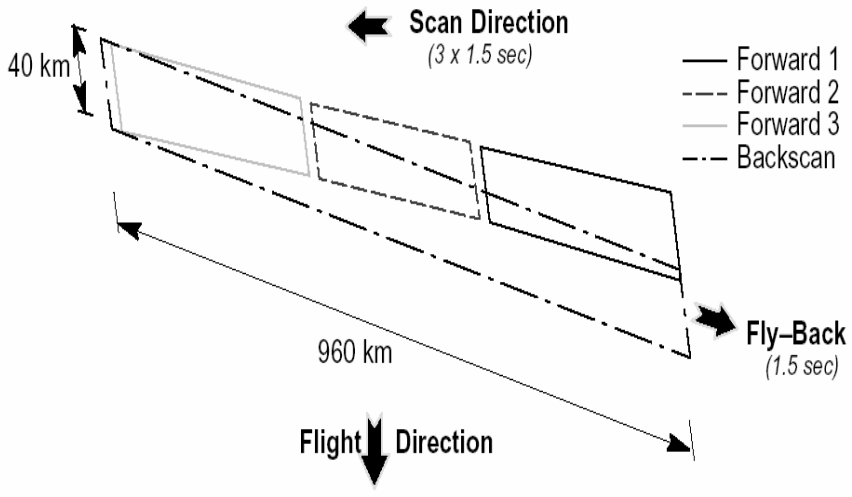


Figure 2.2. GOME scanning processes and its resolution (ERS-2 satellite, http://earth.esa.int/eeo4.96, 1998.).

In our analysis, the GOME data that have cloud fraction > 40% are excluded using an improved GOME cloud product, the GOME Clould AlgoriThm (GOMECAT) (Kurosu *et al*., 1999) because clouds cause the significant UV – VIS signal attenuation. The GOME measurements are affected by the South Atlantic Anomaly (SAA), where radiation is exceptionally strong due to the anomaly of the earth magnetic field (Heirtzler, 2002). The region is to be shown in Figure 2.7. We do not include the region in inverse modeling.

In order to convert slant columns to vertical columns, air mass factors (AMFs) are calculated by a radiative transfer model (LInearized Discrete Ordinate Radiative Transfer model (LIDORT) by Spurr, 2002), which describes the signal through the path of the light in the atmosphere, with the vertical profiles of HCHO taken from the GEOS-Chem simulation (Palmer *et al*., 2001; Martin *et al*., 2002). The AMF is defined as the ratio of the slant column density to the vertical column density,

$$\text{AMF} = (\text{slant column density}) / (\text{vertical column density}) \qquad (2.13)$$

AMF considers the effect of geometry of instrumental and solar viewing angle, and it consider the scattering sensitivity of retrieval species as well. The sources of uncertainty in the AMF calculation are due to the uncertainties in UV albedo, vertical distribution of HCHO, and aerosols. Following the work by Palmer *et al*. (2001) and Martin *et al*. (2002), which formulate the AMF as the integral of relative vertical distribution of the retrieval species (shape factor) over the depth of the atmosphere weighted by altitude-dependent coefficient (scattering weights), we calculate the AMF uncertainties, which are in the range of $1.0 – 13 \times 10^{15}$ molecules cm^{-2}.

The GOME instrument is affected by a diffuser plate problem, and it must be corrected in data retrieval (Martin *et al*., 2002; Palmer *et al*., 2003a). As a result of that artifact, HCHO columns over the remote ocean can be much higher than the background levels due to methane oxidation ($< 5.0\times10^{15}$ molecules cm^{-2}) (Abbot *et al*., 2003). In order to correct the artifact, Martin *et al*. (2002) and Palmer *et al*. (2003a) used the following procedure. First, subtract the mean HCHO column over the Pacific from the corresponding GOME column for each latitude band. Second, add GEOS-Chem simulated HCHO columns over the Pacific to GOME columns. The difference between GOME and simulated HCHO columns over the Pacific reflects the errors due to the diffuser plate artifact.

When applying the same procedure, we find high HCHO variability and noise over the Atlantic, Indian, and Southern Oceans (Figure 2.3). Instead of assuming that the HCHO columns over the Pacific represent background concentrations, we calculate the background HCHO column for each latitude band as the value at the lower 20^{th} percentile. We subtract the background column from GOME HCHO column and then add the 20^{th} percentile value from GEOS-Chem for each latitude band on a daily basis. In addition to the above procedure, we exclude the data that have slant column uncertainties greater than 2σ (σ is the HCHO fitting uncertainty of 4.0×10^{15} molecules cm^{-2}) because these data are usually associated with large HCHO variability over the ocean. We find that the new procedure results in much lower GOME HCHO variability over the oceans, which is shown in Figure 2.3. The magnitude of this correction is about 2×10^{15} molecules

cm^{-2} and the corresponding uncertainty is about 8×10^{14} molecules cm^{-2}. The vertical columns with 20th percentile corrections are about 10 % less than those with the Pacific mean corrections. Taking into account of these uncertainties, we obtain the overall GOME HCHO retrieval uncertainty in the range of 6 – 15×10^{15} molecules cm^{-2} (or 53 – 69% to be shown in Table 2.2).

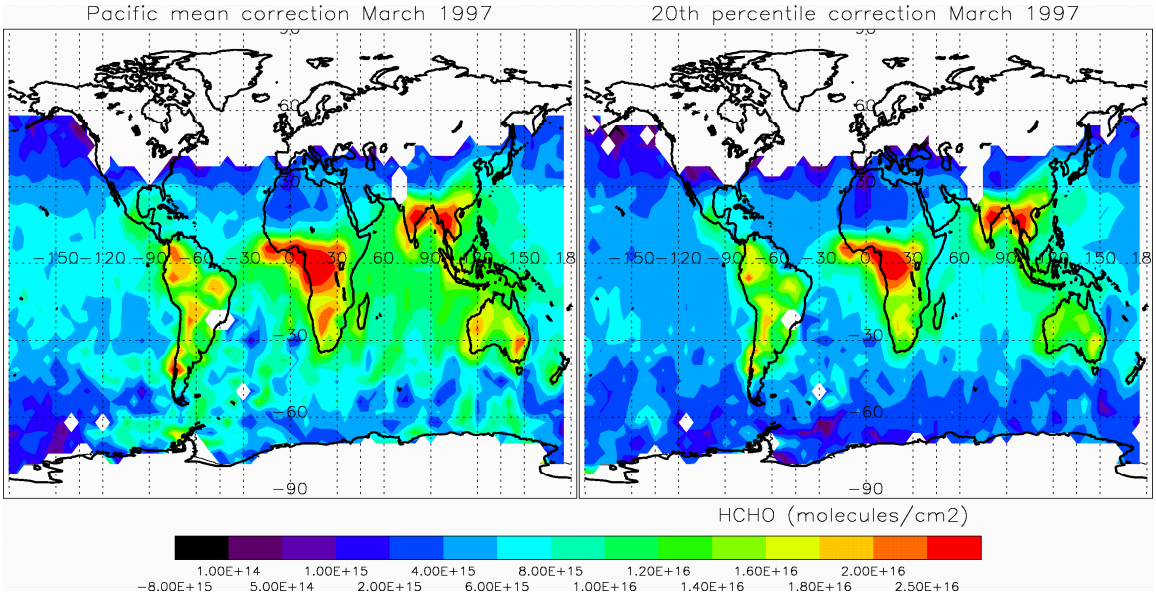


Figure 2.3. Global GOME HCHO vertical columns for March 1997. Left panel shows the result using Pacific mean correction and right panel shows using 20th percentile correction explained in section 2.3. It shows lower and more robust oceanic HCHO columns using 20th percentile corrections.

In inverse modeling, we consider only the regions with relatively high signal-to-noise ratios in GOME measurements. Over these regions, daily GOME HCHO slant columns are > 4σ (1.6×10^{16} molecules cm^{-2}) and the observations satisfying this condition must be

available for more than a season. This criterion is applied for defining the regions for inverse modeling. Shown in Figure 2.4 are such determined source regions over North America (eastern U.S.), Europe (western Europe), East Asia, India, Southeast Asia, South America (Amazon), Africa, and Australia. For the regions at mid latitudes, only during the growing season (May – August) do GOME measurements meet the data selection criterion. As previously stated, the SAA region is not included. The northern equatorial Africa (4 – 12°N) was not included in inverse modeling for reasons described in section 2.5.2.9.

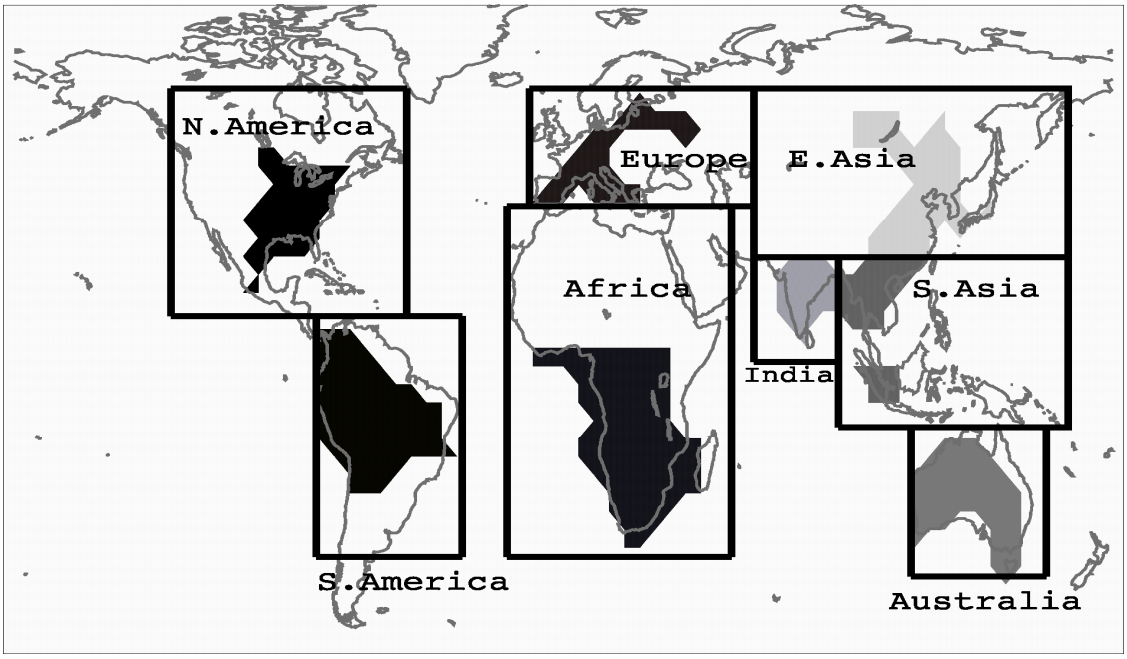


Figure 2.4. Inverse modeling regions with high signal-to-noise ratios in GOME HCHO column measurements are shown by shaded areas. The a posteriori source parameters (state vector) are applied to the rectangle regions in order to estimate the global a posteriori isoprene emissions.

2.4 Model Description and Applications

2.4.1 GEOS-Chem Model

GEOS-Chem is a global 3-D chemical transport model driven by assimilated meteorological data (GEOS-STRAT for 1996 – 1997) from the Global Modeling Assimilation Office (GMAO) (Schubert *et al*., 1993). The 3-D meteorological fields are updated every 6 hours, and the surface fields and mixing depths are updated every 3 hours. We use version 5.05 here with a horizontal resolution of 4°×5° and 26 vertical layers. GEOS-Chem includes a comprehensive tropospheric O_3-NO_x-VOC chemical mechanism (Bey *et al*., 2001), which includes the oxidation mechanisms of 6 VOCs (ethane, propane, lumped $>C_3$ alkanes, lumped $>C_2$ alkenes, isoprene, and terpenes). In the a priori and a posteriori simulations, the model was spun up for a year and the second year's results are used.

Isoprene emissions are estimated using the algorithm of Guenther *et al*. (1995) with updates (Wang *et al*., 1998a; Bey *et al*., 2001). The Global vegetation distribution of 73 ecosystem types is from Olson (1992). A GEOS-Chem grid (4°×5°) contains a maximum of 15 ecosystem types, each of which has its own area fraction and base emission. The seasonality is determined by light intensity, temperature, and LAI. The LAI values are based on satellite observations (Wang *et al*., 1998a). The seasonal cycle of light intensity is largely a function of month. We found that the seasonal cycles of GEOS-STRAT surface temperature in some regions led to disagreement with the observations (section 2.5.2). Bey *et al*. (2001) reduced base isoprene emission rates for several tropical ecosystems (tropical rain forest, tropical montane, tropical seasonal forest, and drought

deciduous) and for grass/shrub by a factor of 3 based on available measurements of isoprene concentrations and fluxes (Helmig *et al*., 1998; Klinger *et al*., 1998). We here use the updates by Bey *et al*. (2001). The resulting global isoprene source during September 1996 – August 1997 is 375 Tg C yr^{-1} (397 Tg C yr^{-1} in 1994 (Bey *et al*., 2001)).

GEOS-Chem uses the emission inventories of industrial NO_X and VOCs for 1985 by Wang *et al*. (1998a). Those emissions are scaled for specific years with national emission data (Bey *et al*., 2001). Biomass burning emissions of CO are constrained by satellite observations of fire counts from the ATSR (Along Track Scanning Radiometer) and AVHRR (Advanced Very high Resolution Radiometer), and AI (Aerosol Index) from TOMS (Total Ozone Mapping Spectrometer) (Duncan *et al*., 2003). Biomass burning emissions of NO_X and VOCs are derived by applying their emission ratios to CO (Wang *et al*., 1998a). We use HCHO/CO molar emission ratios by Andreae and Merlet (2001) as a function of fuel type. The average emission ratio is 0.01. The primary biomass burning emission of HCHO accounts for about 22% of total VOC emissions from biomass burning in the model. The GEOS-Chem simulated HCHO columns are sampled along the GOME orbit tracks at the GOME observation time (10 – 11 a.m. local time) on a daily basis. However, we conduct inverse modeling of monthly averages since GOME HCHO columns have a large day-to-day variability, which is not captured by GEOS-Chem (e.g., Palmer *et al*., 2003a).

2.4.2 Inverse Modeling

The inverse problems in atmospheric tracers are to make statistical inferences about physical parameters of tracer sinks or sources, starting with the measurements of

concentrations (Enting, 2002). Our purpose of inverse modeling is to deduce the better isoprene emission budgets as sources of HCHO with the GOME HCHO measurements. GEOS-Chem 3-D chemical transport model is employed as a forward model with a priori (the information before inversion) parameters regarding to isoprene emissions. The forward model outputs give the atmospheric HCHO.

We apply inverse modeling to estimate the source parameters of HCHO (state vector) using monthly GOME measurements with GEOS-Chem as the forward model. The Bayesian least-squares method is used (Rodgers, 2000). The relationship between the observation vector **y** and state vector **x** can be described as,

$$\mathbf{y} = \mathbf{K}\mathbf{x} + \boldsymbol{\varepsilon} \tag{2.14}$$

where the **K** matrix (Jacobian matrix) represents HCHO sensitivities to the state vector defined by the forward model, and $\boldsymbol{\varepsilon}$ is the error term. There are several types of error that needs to be considered. We consider the measurement error and forward model parameter error. We explained about measurement uncertainties in section 2.3. Forward model parameter error means that of individual state vector in the forward model, which is also important for inverse modeling. We assume that the system is linear and we apply the inversions to minimize the error from fitting the forward model to the observations (Rodgers, 2000).

Minimizing processes which lead to an optimized a posteriori (the information after inversion) value x' is based on Bayesian inference; this theory tells us p(y|x), the conditional probability of the observation given the cause parameters, x (forward

probability density). Based on observations p(x|y), which is conditional probability of the possible sources(cause), given that some effect has been observed. This inverse probability represents our state of knowledge of x after measuring y. With Bayes' theorem,

$$P(x|y) = P(y|x)\, P(x) / P(y) \tag{2.15}$$

P(x) is known as a prior probability (i.e. initial state of knowledge about the parameter), and P(x|y) is the conditional probability function that tells us what we know about the parameter x after the observation (i.e. posterior probability). Basically, the way to change the prior probability into the posterior probability is to multiply the forward probability function usually determined by a forward model (P(y|x)). Least squares in this system can be considered as same as the maximum likelihood with independent uncertainties. Both the observations and the a priori estimates are weighted by estimates of uncertainties.

The uncertainties are used for weightings of the observations and the a priori state vector. We consider the measurement and a priori model parameter errors. The a posteriori state vector $\hat{\mathrm{x}}$ is (Rodgers, 2000),

$$\hat{\mathbf{x}} = \mathbf{x_a} + (\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K} + \mathbf{S_a}^{-1})^{-1}\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}(\mathbf{y} - \mathbf{K}\mathbf{x_a}) \tag{2.16}$$

where $\mathbf{x_a}$ is the a priori state vector, $\mathbf{S_a}$ is the estimated error covariance matrix for $\mathbf{x_a}$, and $\mathbf{S_\varepsilon}$ is the error covariance matrix for observation errors. The a posteriori error covariance matrix is,

$$\hat{\mathbf{S}} = (\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\varepsilon}^{-1}\mathbf{K} + \mathbf{S}_{\mathrm{a}}^{-1})^{-1} \quad (2.17)$$

In inverse modeling, we consider 12 HCHO source parameters in the a priori state vector that contribute to HCHO column concentrations: isoprene emissions from 9 different vegetation groups, other biogenic VOCs including isoprene emissions from the rest of vegetation groups, and two additional HCHO sources from biomass burning and industry. Table 2.1 shows the global isoprene emissions estimates by Guenther *et al*. (1995) for the defined vegetation groups. Some vegetation groups are the major isoprene emitters for specific regions: agricultural lands for India, dry evergreen and crop/woods for Australia, and regrowing woods for North America. Figure 2.5 shows the global distribution of the vegetation groups.

Table 2.1. Isoprene emitting ecosystem groups applied in inverse modeling.

Ecosystem groups	Olson Code	Global emissions (Tg C yr^{-1})	A priori uncertainty (%)
1. Tropical rain forest	33	84.4	300
2. Grass /shrub (hot, cool)	40, 41	91.7	400
3. Savanna	43	48.3	400
4. Tropical seasonal forest & thorn woods	29, 59	80.1	300
5. Temperate mixed & temperate deciduous	24, 26	11.3	300
6. Agricultural lands	31, 36	20.9	300
7. Dry evergreen & crop/woods (warm)	48, 58	16.9	300
8. Regrowing woods	56	27.5	300
9. Drought deciduous	32	60.5	300
10. Rest of the ecosystems[1]		61.4	300
		Total 503	

Ecosystem types are defined by Olson (1992). The global isoprene emissions are taken from Guenther *et al.* (1995). [1] Includes all other ecosystems with biogenic emissions assigned by Guenther *et al.* (1995).

The measurement errors from GOME retrievals are discussed in section 2.3. As discussed in the section 2.2.2, transport does not significantly affect GOME HCHO measurements used in inverse modeling, which are dominated by isoprene. Transport error is therefore not included in inverse modeling. The model parameter errors reflect the uncertainties of the source parameters in the forward model. We assign the source parameter errors for the vegetation types with field or laboratory measurements to 300% (Guenther *et al*., 1995) and to 400% for those without measurements. The assigned isoprene emission errors include all the variables in the model for estimating the natural VOC emissions (Guenther *et al*., 1995).

The emissions of HCHO from biomass burning are calculated using the CO biomass burning emission inventory and the HCHO/CO molar emission ratios. The CO biomass burning emissions have an uncertainty of 50% (Palmer *et al*., 2003b), while the HCHO/CO molar emission ratios as a function of fuel type were based on limited measurements compiled by Andreae and Merlet (2001). Therefore, we assume the uncertainty of HCHO emission from biomass burning is as high as that of isoprene (300%). We assign the error of 50% for industrial emissions (Palmer *et al*., 2003b). As stated in section 2.2.2, CH_4 provides the background source of HCHO. We do not include the CH_4 contributions in the inverse modeling because the uncertainty of HCHO from CH_4 oxidation is much smaller than that of isoprene emissions. Assigning the relatively small uncertainty to HCHO produced from CH_4 oxidation in inverse modeling results in no a posteriori change in this source.

In the forward model calculations, we compute the sensitivity of HCHO columns to the emissions from the 12 source categories. The sensitivity calculation of HCHO

columns to each source category is complicated by the feedbacks of these emissions on OH concentrations. When reducing isoprene emissions, OH concentrations increase affecting HCHO production and loss. In our calculations, we archive hourly OH, NO, and O_3 concentrations from the standard simulation. The sensitivity of HCHO columns to each source category is calculated by removing that source while holding hourly OH, NO, and O_3 concentrations to the values in the standard model. The procedure is necessary because the inverse model assumes that the Jacobian matrix is linear. The validity of the linear assumption is then checked by conducting a full chemistry simulation with the a posteriori sources and comparing the resulting HCHO columns with the linear projections by the inverse model (section 2.5.3). In this study, we find that the sensitivities of source parameters are close to linear in the range of emission variability for all 8 regions with high signal-to-noise ratios in GOME HCHO measurements.

We apply the inverse modeling to each region separately because the same vegetation types on different continents in the Olson (1992) map can have different species compositions (A. Guenther, personal communication, 2003). The regions for inverse modeling are shown in Figure 2.4. The selection of the significant source parameters in the state vector to avoid numerical errors in the inversion is described in section 2.5.3. In order to estimate the global a posteriori isoprene inventory, we extend a posteriori source parameters for isoprene for the 8 regions (shaded regions in Figure 2.4) to their respective continents between 40°S – 60°N where > 95% of isoprene emissions take place (rectangle regions in Figure 2.4) by scaling each ecosystem emission with the a posteriori / a priori ratio of the corresponding source parameter (section 2.6).

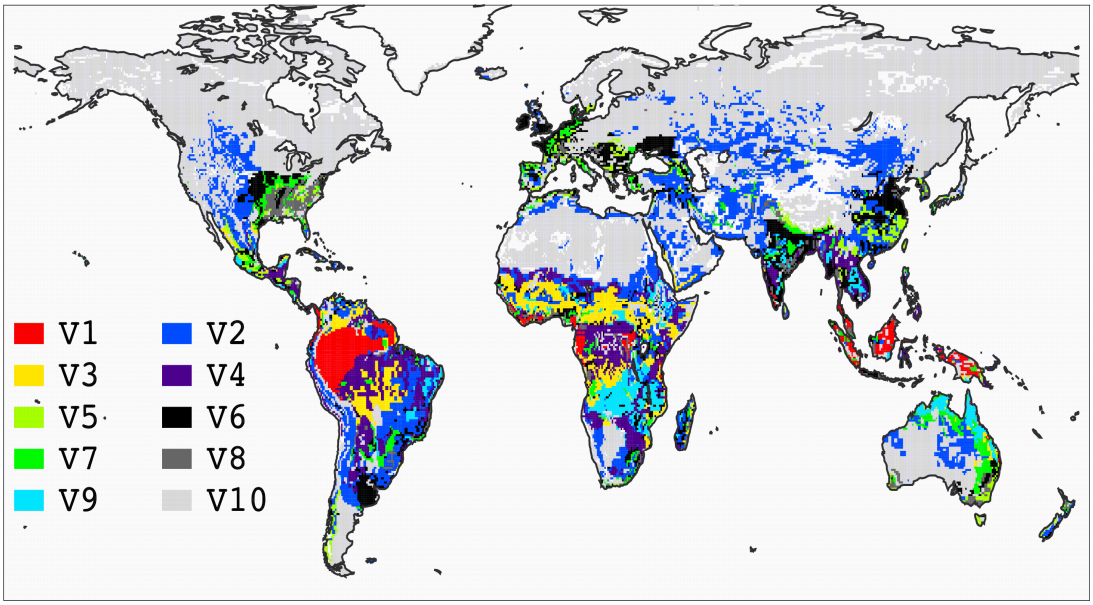


Figure 2.5. Global distribution of the 10 ecosystem groups (Table 2.1) applied in inverse modeling, including tropical rain forest (V1), grass/shrub (V2), savanna (V3), tropical seasonal forest & thorn woods (V4), temperate mixed & temperate deciduous (V5), agricultural lands (V6), dry evergreen & crop/woods (warm) (V7), regrowing woods (V8), drought deciduous (V9), and the rest of ecosystems (V10). The ecosystem types are defined by Olson (1992) with a resolution of 0.5°×0.5°.

2.5 Analysis

The annual global distributions of isoprene for the GEIA 1990 inventory (Guenther *et al*., 1995), and the September 1996 – August 1997 a priori and a posteriori GEOS-Chem estimates are presented in Figure 2.6. The a posteriori isoprene emissions are increased (vs. the a priori emissions) over all the regions, particularly over the tropics. Compared to the GEIA inventory, the a posteriori isoprene emissions are generally higher at mid latitudes but lower in the tropics. The exception is the much higher a posteriori estimate over Australia. Table 2.2 shows the estimates of the annual isoprene emissions from the 8 regions.

2.5.1 Observed and Simulated HCHO Columns

The observed annual mean global HCHO columns from GOME and the corresponding columns simulated GEOS-Chem with the a priori and a posteriori emissions are shown in Figure 2.7. The a priori GEOS-Chem simulation significantly underestimates the global HCHO columns by > 30%. The mean correlation coefficient between the a priori model simulation and GOME observations is 0.68. The discrepancies are particularly large over the tropical South America and Africa, where more than one third of global isoprene is emitted according to the a priori isoprene emissions.

Table 2.2. Regional statistics of GOME and simulated HCHO columns, and the priori, posteriori and GEIA estimates of the annual isoprene emissions for the inversion regions[1].

	GOME	Weighted uncertainties[2](%)		Correlation coefficient(R)[3]		Model bias (%)		Isoprene emission (Tg C/yr)		
Regions	Ω (%)[4]	priori	posteriori	Priori	Posterior i	priori	posterior i	Priori	posteriori	GEIA
North America	59	291	69	0.84	0.84	-14.3	-3.6	22.2	25.7	21.4
Europe	69	287	96	0.52	0.60	-29.9	-11.9	9.5	12.0	6.1
East Asia	56	280	63	0.63	0.75	-39.2	-18.6	17.4	24.8	12.8
India	59	285	122	0.57	0.56	-33.2	-18.4	10.5	14.4	15.2
Southeast Asia	54	298	110	0.66	0.69	-35.8	-19.4	20.2	29.1	38.2
South America	54	337	75	0.58	0.64	-31.8	-12.6	79.4	106.4	163.5
Africa	53	332	102	0.56	0.54	-46.3	-23.6	60.3	103.3	105.7
Australia	69	302	96	0.52	0.56	-40	-24.8	33.3	50.6	31.1
Global	60			0.68		-35		375	566	503

[1] The values are for the shaded area (Figure 2.4) during the growing seasons.
[2] Weighted uncertainties of the state vector (source parameters).
[3] Calculated based on 4°×5° monthly mean GOME and model data.
[4] Ω denotes the overall percentage GOME retrieval uncertainties with respect to the vertical columns.

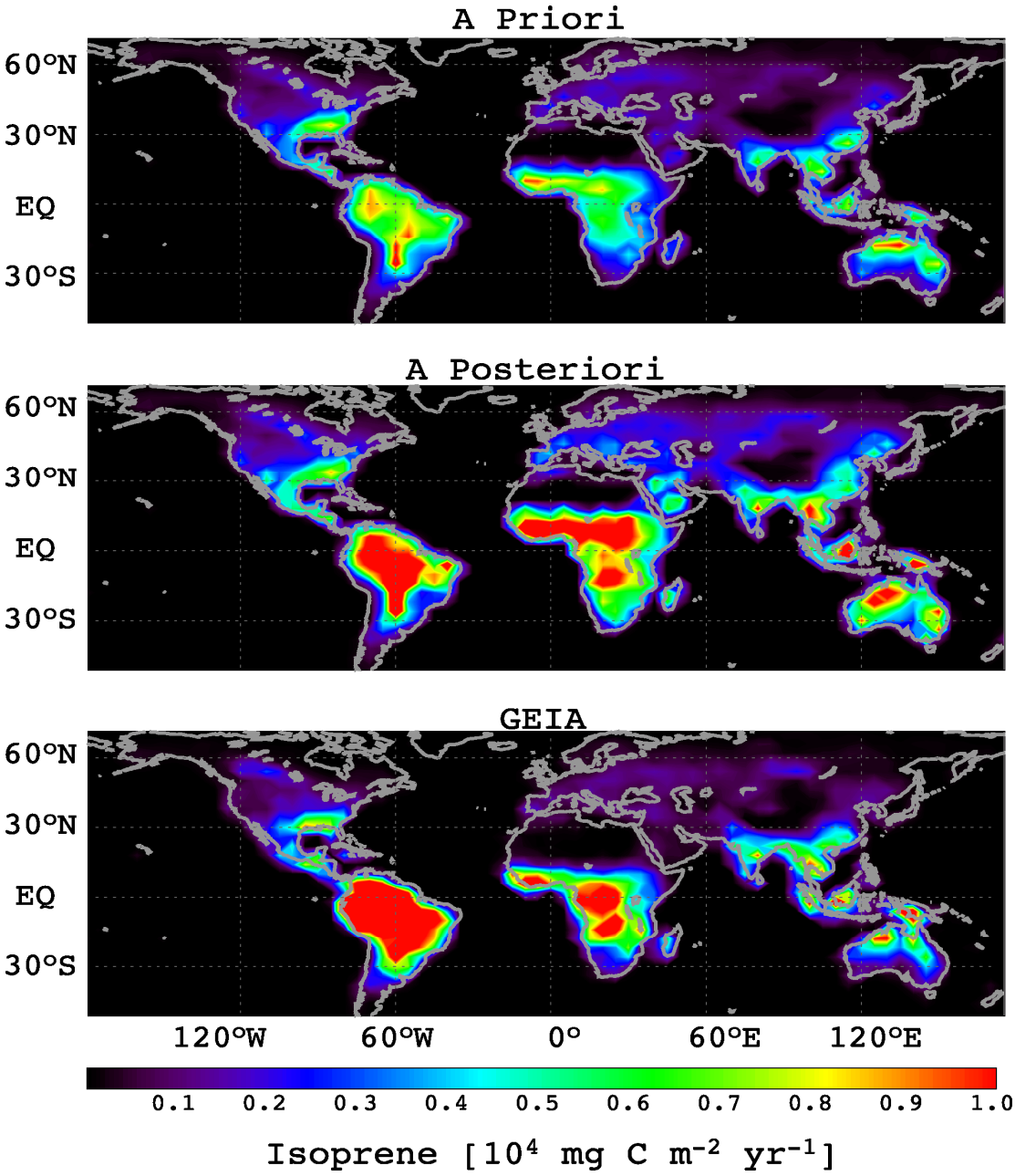


Figure 2.6. Estimated annual global distributions of isoprene emissions [10^4 mg C m^{-2} yr^{-1}]. Upper: The GEOS-Chem simulation with the a priori isoprene emissions for September 1996 – August 1997. Middle: Same as the upper panel but with the a posteriori isoprene emissions. Bottom: The GEIA inventory for 1990 (Guenther *et al.*, 1995).

The uncertainties of GOME retrievals are large in the range of 53 – 69% over the 8 selected regions, even though the regions are chosen based on high signal-to-noise ratios. Table 2.2 shows the GOME uncertainty, correlation coefficient, and model bias for each region of inverse modeling. North America has the smallest bias (-14%) during the growing season (May – August). Africa has the highest bias (-46%).

Figure 2.7 shows the annual mean GEOS-Chem HCHO columns with a posteriori emissions. Full model simulations using the a posteriori emissions are in much better agreement with GOME observations over high isoprene regions. The regional biases of the a priori model are reduced by about 50% (Table 2.2). The a posteriori uncertainties are greatly reduced due to the constraints by GOME observations. The a posteriori model still has a low bias compared to GOME observations partly because of the relatively large GOME measurement uncertainties. Despite these improvements, there is still serious disagreement between GOME and a posteriori HCHO columns over the northern equatorial Africa (Figure 2.7), which will be discussed in section 2.5.2.9.

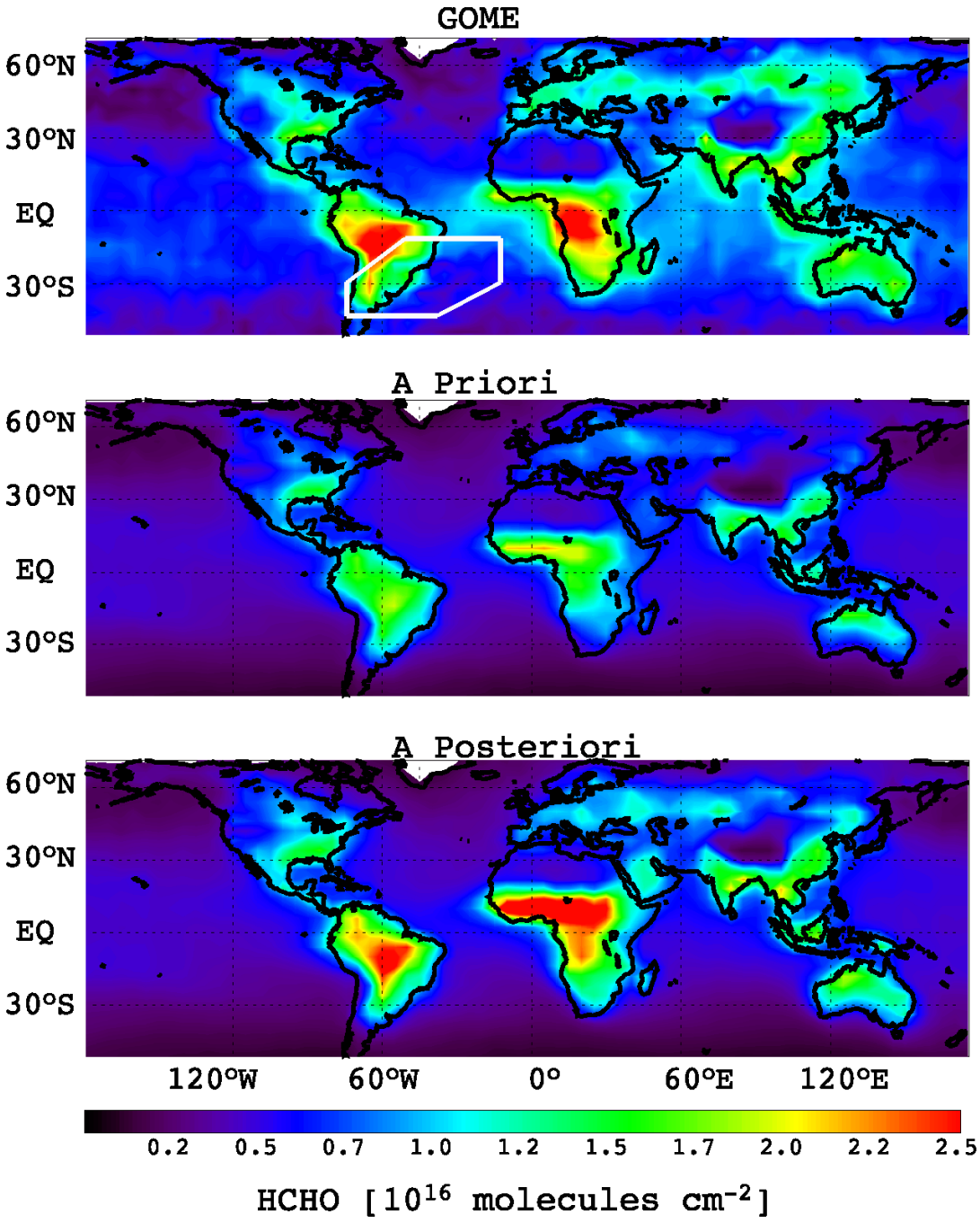


Figure 2.7. Annual mean observed and simulated vertical HCHO columns for September 1996 – August 1997. Upper: GOME retrieved columns. Middle: The a priori GEOS-Chem columns. Bottom: The a posteriori GEOS-Chem columns. The GEOS-Chem HCHO columns shown are coincident in space and time with GOME measurements. The white polygon shows the region of the South Atlantic Anomaly.

2.5.2. Regional Isoprene Emissions

Figure 2.8 compares simulated and observed GOME monthly column concentrations of HCHO over the 8 selected regions (shaded regions in Figure 2.4). Table 2.3 shows the contribution by each emission category to both a priori GEOS-Chem and a posteriori inverse-model projected HCHO columns. The relative a posteriori changes of HCHO contributions from different vegetation groups are the same as those from isoprene emissions since the inversion is linear. We discuss the results by region in the following sections.

2.5.2.1 North America (Eastern U.S.)

We compare observed and simulated monthly HCHO columns (4°×5°) only for the growing season between May – August (section 2.3). The corresponding data correlation between a priori estimates and GOME observations is rather high. This correlation coefficient (R) of 0.84 is comparable to the previous study (for July 1996) (Palmer *et al.*, 2003a). According to the a priori estimate, regrowing woods is the largest isoprene emission group, and temperate mixed and temperate deciduous are the second largest group. The monthly contributions of the major a priori sources to HCHO columns are shown in Figure 2.9. The oxidation of CH_4 provides the background levels of HCHO columns (~25% of the total). The a priori biogenic emissions account for 63% of HCHO columns.

The a posteriori source parameters suggest relatively small changes for most emission categories (Table 2.3), which implies that isoprene emissions over North America are relatively well estimated. The result is expected because the base emissions for vegetations in this region and Europe are better measured than over the other regions

(Guenther *et al*., 1995). The a posteriori biogenic emissions of 26 Tg C yr^{-1} in this region are about 15% higher and the resulting HCHO columns are in better agreement with GOME.

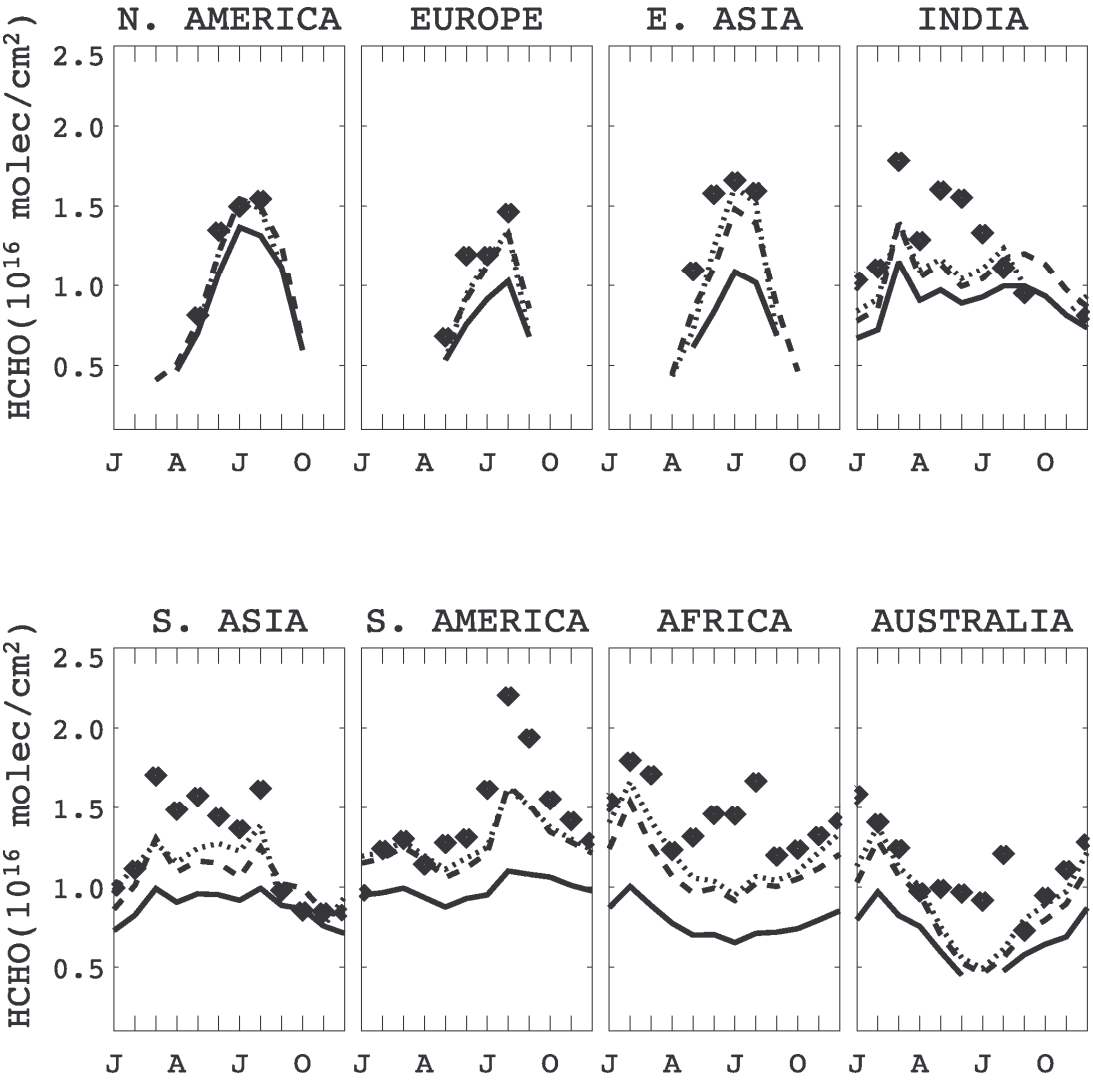


Figure 2.8. Monthly mean HCHO column concentrations in the 8 regions (Fig. 2.4) during September 1996 – August 1997. The time sequence is reordered to January through December. The diamonds show GOME column concentrations. The solid lines show the corresponding GEOS-Chem simulated columns with the a priori sources. The dashed lines are GEOS-Chem simulated columns with the a posteriori sources. The dotted lines show the linearly inverse-model-projected HCHO columns with the a posteriori sources. The values below the GOME detection limit (4.0×10^{15} molecules/cm^2) are not shown.

2.5.2.2. Europe

This region has the a priori isoprene emissions of 9.5 Tg C yr^{-1}. As for North America, we consider only the growing season between May and August. The a priori biogenic emissions account for 58% of HCHO columns in that season (Table 2.3). The largest change in the a posteriori emissions is in those from agricultural lands. The inverse modeling results suggest an increase of this source by a factor of 3.4 (Table 2.3). Similarly large increases for this source category are found for East and Southeast Asia. It is possible that this problem is due to incorrect classification of vegetation types. Another possibility is that crop and farming practice over those regions are different from North America, resulting in different emissions. The total a posteriori isoprene source of 12 Tg C yr^{-1} for the region is 26% higher than a priori (Table 2.2). The model bias improves from the a priori –29.9% to a posteriori –11.9%.

2.5.2.3. East Asia

This region includes eastern China, Korea, and Japan with the a priori isoprene source of 17 Tg C yr^{-1}. The seasonal variation is similar to that of other northern mid latitude regions. We consider only the growing season (May – August) for inverse modeling. The a priori biogenic emissions account for 56% of HCHO columns in that season. Although the total isoprene source change is 43% (Table 2.2), large changes are found for individual sources including a factor of 2 – 4 increase in the emissions from grass/shrub, agricultural lands, and regrowing woods. The increase of the other biogenic sources is also significant. The emissions from mixed deciduous forests decrease by a factor of 10 (Table 2.3). These large changes appear to indicate problems in the Olson (1992) vegetation distribution. The a posteriori biomass burning source of HCHO increases by a

factor of 4 (Table 2.3), making it the most significant source in spring (not shown). After an increase of 43% in the a posteriori isoprene emissions to 24.8 Tg C yr^{-1}, the model is still biased low by 18.6% (Table 2.2).

2.5.2.4. India

The rapid increase of GOME HCHO columns in March (Figure 2.8) coincides with biomass burning in this region. The model captures this seasonal change. However, the simulated biogenic emissions do not show the large increase from winter to summer as observed. Therefore, the a posteriori model underestimates GOME HCHO columns in summer. The sources from industry and CH_4 oxidation do not have a large seasonal variation either. The most likely candidate to explain the seasonal change of GOME HCHO column is the biogenic sources. The current emission algorithm apparently does not simulate the seasonal cycle correctly. This region has the a priori isoprene emissions of 11 Tg C yr^{-1}. Inverse modeling suggests a factor 2 – 3 increase for isoprene emissions from grass/shrub, regrowing woods, and the rest of vegetation category. With an increase of 37% to 14.4 Tg C yr^{-1}, the model bias improves from the a priori –33.2% to a posteriori –18.4%.

2.5.2.5. Southeast Asia

This region includes the Indochina Peninsula and Indonesia (4°S – 30°N) with the a priori isoprene source of 20 Tg C yr^{-1}. The biogenic emissions account for 40% of HCHO column concentrations. There are two features in the region that are similar to India. First, the large March maximum (Figure 2.8) is largely due to biomass burning emissions. The inverse modeling suggests a factor of 2 increase of this source. Second, the observed large seasonal shift from winter to summer cannot be reproduced by the model. Large

increases of isoprene emissions (by about a factor of 3) are needed for agriculture lands and the rest of vegetations group. The a posteriori isoprene emissions increase by 44% to 29.1 Tg C $yr^{-1}$and the model bias is reduced from the a priori -35.8% to a posteriori -19.4%.

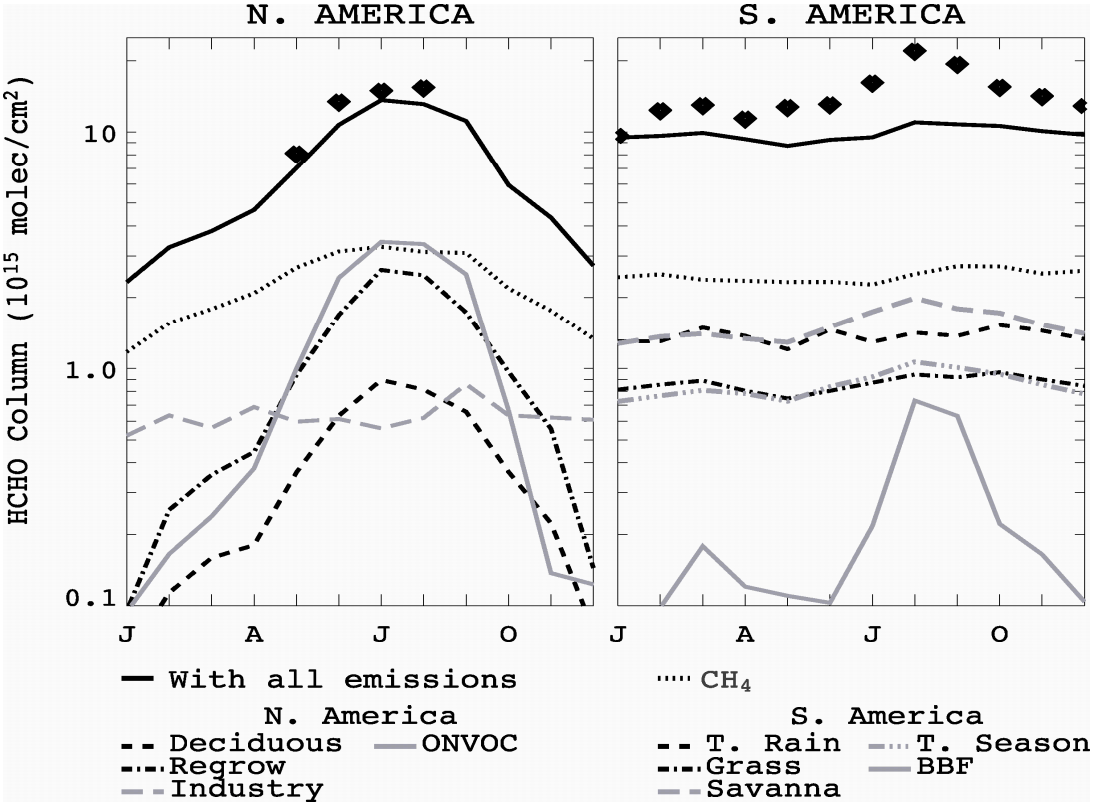


Figure 2.9. Contributions of the a priori sources to the simulated monthly mean HCHO column concentrations over North America (Eastern U.S.) and South America (Amazon). The diamonds are GOME HCHO columns. "With all emissions" denotes the simulated HCHO column concentration with all emission sources. "CH_4" denotes HCHO from CH_4 oxidation. The other source contributions are: "Deciduous" (temperate mixed and temperate deciduous), "ONVOC" (isoprene from the other ecosystems), "T. Rain" (tropical rain forest), "T. Season" (tropical seasonal forest and thorn woods), "Regrow" (regrowing woods), "Grass" (grass/shrub), and "BBF" (biomass and biofuel burning).

2.5.2.6. South America (Amazon)

This region including Amazon has a large source of isoprene (~80 Tg C yr^{-1}), accounting for > 20% of the global a priori source. Biogenic emissions account for 65% of the simulated HCHO column concentrations here. GOME HCHO observations show a distinct seasonal variation with a maximum in July – September (Figure 2.8). The simulated HCHO monthly columns are lower with a smaller seasonal variation. Figure 2.9 shows the source contributions of isoprene from tropical rain forest, tropical seasonal forest, grass/shrub, savanna, and that from biomass burning. Among the major sources, the contributions from biomass burning, savanna, and tropical seasonal forest emissions show the observed austral spring maximum.

To capture the observed seasonal variation, the inverse modeling results suggest a factor of 5 increase in biomass burning, a factor of 2 increase in tropical seasonal forest emissions and a factor of 2 decrease in grass/shrub emissions. The posteriori isoprene emissions increase by 34% to 106 Tg C yr^{-1}, but are still 35% less than the corresponding GEIA estimate. The a posteriori model bias is decreased to -12.6%.

2.5.2.7. Africa

This region (40°S – 4°N) has the a priori isoprene emissions of 60 Tg yr^{-1}. Biogenic emissions account for > 50% of the model HCHO columns. The a priori simulation shows a large underestimate of the GOME observations (model bias: -46%). The inverse modeling results suggest large increases (a factor of 2 – 4) in emissions from tropical rain forest, grass/shrub, drought deciduous, and the other vegetation group. The HCHO source from biomass burning is increased by factor of 2 (Table 2.3). The a posteriori model reproduces reasonably well the seasonal increase from austral spring to summer due to

increasing biogenic emissions, but fails to capture the high values in June-August. There are significant discrepancies between the simulated and GOME HCHO columns over the northern equatorial Africa (Figure 2.8), where the model overestimates the observations. We investigate the causes in section 5.2.9. The inverse modeling shows a 71% increase of isoprene emissions 103.3 Tg C yr^{-1} and the a posteriori model bias is decreased by 49% to -23.6%.

2.5.2.8. Australia

This region (12 – 40°S) shows a seasonal cycle typical for the southern hemisphere (Figure 2.8). Biogenic emissions account for 70% of the model HCHO columns. The a priori simulation greatly underestimates GOME observations in this region (-40%). The inverse modeling results suggest significant increases in the emissions from savanna, regrowing woods, and the rest of vegetations group. With a 52% increase of isoprene emissions to 51 Tg C yr^{-1}, the a posteriori model bias is decreased to -24.8% (Table 2.3).

The model shows a small minimum in January when GOME observations show a maximum. The simulated decrease is due to a corresponding change in GEOS-STRAT surface temperature, which reduces isoprene emissions (not shown). A more prominent illustration of a similar problem over the northern equatorial African region is discussed in the next section. The European Centre for Medium-range Weather Forecasting (ECMWF) surface temperature for the same period does not show this seasonal change, likely indicating a problem in GEOS-STRAT surface temperature simulation for this region.

Table 2.3. Mean source contributions to the a priori and inverse model-projected HCHO columns (10^{14} molecules cm^{-2})[1].

	N. America		Europe		E. Asia		India		S. Asia		S. America		Africa		Australia	
	Pri	pos	Pri	Pos	Pri	Pos	Pri	Pos	Pri	pos	Pri	Pos	pri	Pos	pri	pos
V1	-	-	-	-	-	-	**-**	**1.5**	**5**	**7.9**	**13.8**	**17.9**	**2**	**6.4**	-	-
V2	**6.1**[2]	**8.5**	**2.3**	**5.4**	**8.1**	**21.8**	**2.5**	**6.5**	1.6	1.6	**8.6**	**4.3**	**2.7**	**10.3**	**4.8**	**2.9**
V3	1	1	-	-	-	-	-	-	-	-	**15.6**	**21.4**	**10.5**	**14.6**	**1.2**	**5**
V4	-	-	-	-	-	-	-	-	**4.6**	**6.9**	**8.5**	**16.2**	**4.5**	**4.5**	-	-
V5	**7.8**	**13.3**	**5.7**	**10.3**	**12.5**	**1.3**	-	-	**3.2**	**4.8**	-	-	-	-	**1.6**	**1.6**
V6	**5.6**	**8.9**	**4.8**	**16.3**	**4.9**	**12.8**	**8.3**	**7.4**	**4.9**	**14.3**	1.2	1.2	2.1	2.1	**-**	**1.8**
V7	**4.6**	**3.2**	**1.2**	**2**	-	-	**1.9**	**1.9**	1.2	1.2	**2.6**	**2.6**	1.4	1.4	**14.5**	**10.1**
V8	**22.6**	**18.1**	**3.7**	**4.1**	**1.3**	**5.2**	**7.5**	**13.5**	**8.3**	**8.3**	1.6	1.6	**3.4**	**3.4**	**-**	**5.9**
V9	-	-	-	-	-	-	**2.4**	**3.4**	**3.5**	**1.7**	**1.8**	**1.6**	**6.9**	**15.2**	**4.3**	**4.7**
RV	**30.7**	**40**	**34.6**	**38**	**28.5**	**48.5**	**3.4**	**8.3**	**3.7**	**11.7**	**9.9**	**17.9**	**8.4**	**21.8**	**16.2**	**35.9**
BB	-	-	1.8	1.8	**3.4**	**14**	**9.9**	**10.8**	**5.6**	**11.3**	**2.3**	**11.3**	**4.2**	**8.8**	-	-
IND	**6**	**7.2**	**5.4**	**6.4**	**4.8**	**8.1**	**5.2**	**7.7**	**4.2**	**5.1**	1	1	1.2	1.2	-	-
Total	125	141	90	115	98	146	88	108	89	118	98	129	78	121	67	91

[1]The inversion is applied for the individual region. All the inversion quantities here are for the shaded areas in Figure 1. We considered only the growing season (May – August) for the regions at mid latitudes. Ecosystem classification is from Olson (1992). Inverse model-projected HCHO columns are the products of the ratio of the a posteriori / priori source parameters and the corresponding a priori HCHO columns.

[2]Bold faced values denote that the vegetation types are included in the state vector. The number of state vectors for each region is listed in Table 2.4.

"- "denotes the values < 1.0 x 10^{14} molec/cm^2.

"pri": The a priori emission contributions.

"post": the a posteriori inverse-model projections.

"V1 – V9": Isoprene contributions from tropical rain forest (V1), grass/shrub (V2), savanna (V3), tropical seasonal forest (V4), temperate mixed & temperate deciduous (V5), agricultural lands (V6), dry evergreen and crop/woods (V7), regrowing woods (V8), drought deciduous (V9), the rest of biogenic sources (RV), biomass and biofuel burning (BB), and industrial VOC emissions (IND). "Total" also includes CH_4 oxidation.

2.5.2.9. Discrepancies over the Northern Equatorial Africa

The largest discrepancy in the seasonal HCHO column variations between GOME observations and GEOS-Chem simulations is found over the northern equatorial Africa (4 – 12°N). The dominant ecosystems over this region are savanna, tropical seasonal forest & thorn woods, drought deciduous, and grass/shrubs. The distributions and ecosystem

types are different from other African regions (Figure 2.8). This region is excluded in the inversion for Africa due to the large discrepancy between observed and simulated HCHO seasonal cycles.

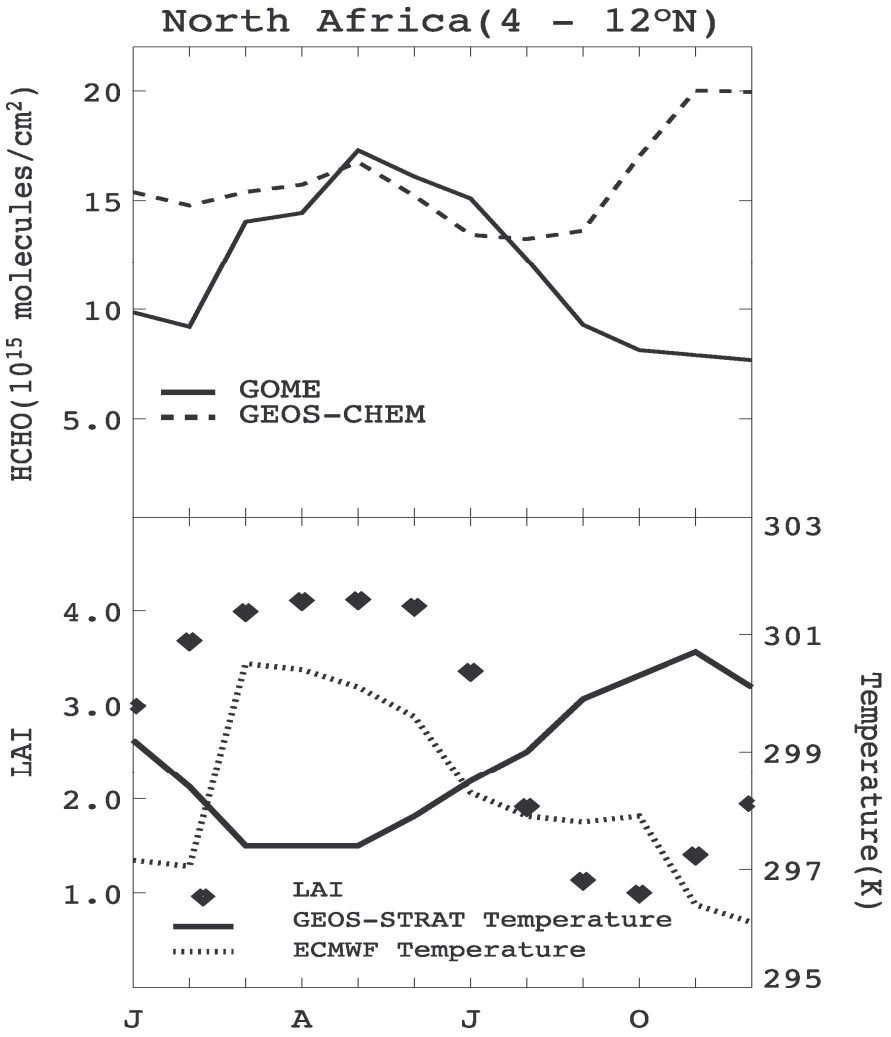


Figure 2.10. The discrepancy between monthly GOME measured and GEOS-Chem simulated HCHO columns over the northern equatorial Africa (4 – 12°N). The corresponding monthly mean LAI, ECMWF surface temperature, and GEOS-STRAT surface temperature are shown in the lower panel. There are no GOME HCHO measurements that match our data selection criteria for inversion in November 1996 over this region.

The GEOS-Chem monthly mean HCHO columns show a seasonal maximum in November and December, whereas GOME shows a maximum in May (Figure 2.10). The monthly variations of the GEOS-STRAT and ECMWF surface temperature and GEOS-Chem LAI for this region are also shown in Figure 2.10. The seasonal cycle of LAI is consistent with that of monthly GOME HCHO columns. However, the GEOS-STRAT surface temperature has an opposite seasonal cycle. In comparison, the seasonal cycle of the corresponding surface temperature simulated by ECMWF is consistent with GOME observations. Everything else being the same, isoprene emissions increase by 50% when temperature increases from 297 to 301K.

The GEIA emission inventory does not have the bias likely because the International Institute for Applied Systems Analysis (IIASA) monthly mean climatological surface temperature field (Leemans *et al*., 1992) was used. For the same reason, Wang *et al*. (1998a) showed much lower isoprene emissions in January than July over the northern equatorial Africa. The discrepancy appears to be caused by GEOS-STRAT overestimates of surface temperature for this region in fall and winter.

2.5.3. Degrees of Freedom of the Jacobian Matrix and Nonlinearity

The state vector has a total of 12 source parameters in the inverse model. We selected only the parameters with significant contributions based on the a priori emissions because the noises introduced by small sources are sometimes manifested in the a posteriori results for significant sources. The resulting significant source parameters in the inversion are 7 to 9 for each region (Table 2.4). It is important to know if GOME HCHO measurements provide enough information to resolve these parameters. We evaluate the degree of freedom in inverse modeling by calculating the singular values of the pre-

whitened Jacobian matrix ($\overline{K} = S_{\varepsilon}^{-1/2} K S_a^{1/2}$) (Rodgers, 2000; Heald *et al*., 2004); the degree of freedom is defined by the number of singular values > 1. We find that GOME observations are generally enough to resolve the significant parameters in inverse modeling (Table 2.4) because the estimated number of significant parameters is close to the state vector size. The slightly large size of the state vector indicates some inter-dependence in the a posteriori estimates of source parameters.

The sensitivities of the source parameters to HCHO columns in the inversion are assumed to be linear. We compare the linear projection of the inverse modeling with the a posteriori full chemistry simulation in Table 2.4. The largest nonlinearity of ~8% is over Southeast Asia and Africa. The smallest nonlinearity of < 2% is over North America and Europe. The other regions have a nonlinearity of 2 – 5%.

Table 2.4. Samples, state vector size, significant eigenvalues, and nonlinearity

	N. America	Europe	E. Asia	India	S. Asia	S. America	Africa	Australia
Samples[1]	152	148	216	162	261	660	792	564
State vector size[2]	7	7	7	9	9	8	8	8
Significant eigenvalues[3]	6	5	6	6	8	7	7	6
Nonlinearity (%)[4]	1.7	0.4	5.3	4.3	8	2.1	7.8	7.5

[1] The number of monthly mean GOME HCHO measurements that meet our criteria for the usage of inversion.
[2] The number of significant parameters (see text for details)
[3] The number of singular values of the pre-whitened Jacobian that are >1.
[4] [1 – (the a posteriori simulated HCHO column) / (inverse-model linearly projected HCHO column)]×100.

2.6 Effects of the Isoprene Emission Change on Global OH and O_3 Concentrations.

We extended the inverse modeling results (Table 2.3) to the adjacent continental regions (Figure 2.4) to estimate the isoprene emissions for each continent. We estimate the a posteriori global isoprene source of 566 Tg C yr^{-1}. It is ~50% higher than that of the a priori source (375 Tg C yr^{-1}), but is slightly higher (12%) than the GEIA inventory of 503 Tg C yr^{-1} (Guenther *et al*., 1995). The a posteriori global isoprene source is also comparable to that of 597 Tg C yr^{-1} estimated by Wang *et al*. (1998a). The a posteriori isoprene emissions from tropical rain forest and tropical seasonal forest are increased by about 60% globally.

When compared to the GEIA inventory, these tropical emissions are lower (globally) by about 30% (Table 2.5). Therefore GOME HCHO observations support a general reduction of isoprene emissions from these tropical ecosystems but not as drastic as assumed in the a priori model. Of particular interest is that a posteriori isoprene emissions suggest that the reduction depends strongly on the continent. For example, the a posteriori reduction for the tropical rain forest emissions is a factor of 2.5 from GEIA over South America but is negligible over Africa (Table 2.5).

Isoprene is the most dominant reactive VOC in the troposphere. The increase of isoprene emissions affects global tropospheric OH and O_3 concentrations (Wang and Jacob, 1998; Spivakovsky *et al*., 2000). Generally the OH concentrations are reduced by reacting with isoprene and its oxidized products such as methyl vinyl ketone. The tropospheric annual mean OH concentration calculated by the a posteriori GEOS-Chem simulation decreases by 10.8% to 0.95×10^6 molecules/cm^3 resulting in an atmospheric methyl chloroform (CH_3CCl_3) lifetime against the tropospheric OH of 5.7 years (5.2

years for the a priori simulation), in better agreement with 5.99 (+0.95/-0.71) years estimated by Prinn *et al*. (2001). The atmospheric CH_3CCl_3 lifetime is estimated in the same manner as Bey *et al*. (2001).

The percent decreases of annual and zonal mean concentrations of OH and NO_x due to the increase of the a posteriori isoprene and the other biogenic emissions are shown in Figure 2.11. The upper tropospheric reduction of OH in the tropics is due in part to convective transport of isoprene and its oxidation products to the region and in part to the significantly reduced NO_x concentrations. The loss of NO_x is due to the formation of PAN through the reaction of peroxyacetyl radicals and NO_2. Isoprene is a major precursor for peroxyacetyl radicals. The lifetime of PAN strongly depends on temperature. The effect is most significant in the upper troposphere, where low temperature leads to a long lifetime of PAN.

Table 2.5. Ratios of the a posteriori isoprene base emission rates to those of GEIA.

	N. America	Europe	E. Asia	India	S. Asia	S. America	Africa	Australia
V1	-	-	-	1.26	0.48	0.39	0.96	-
V2	0.42	0.69	0.81	0.78	-	0.15	1.14	0.18
V3	-	-	-	-	-	1.12	1.12	3.44
V4	-	-	-	-	0.59	0.74	0.39	-
V5	1.19	1.26	0.07	-	1.05	-	-	0.70
V6	1.28	2.08	2.08	0.72	2.32	-	-	3.28
V7	0.63	1.44	-	0.9	-	0.90	-	0.63
V8	0.64	0.88	3.20	1.44	0.80	-	0.80	5.20
V9	-	-	-	0.42	0.15	0.27	0.66	0.33
RV	1.04	0.88	1.36	1.92	2.56	1.44	2.08	1.76

The definitions of vegetation types are listed in Table 2.3. Only vegetation groups included in the state vector are shown.

The decrease of OH is also notable in the lower and upper troposphere at northern midlatitudes (30°– 60°N). It is due to isoprene emissions over North America, Europe, and Siberia in summer. The impact of increased isoprene emissions on global tropospheric O_3 is much less than that of OH. The a posteriori tropospheric annual mean global O_3 burden increases only by 1.5% to 333 Tg.

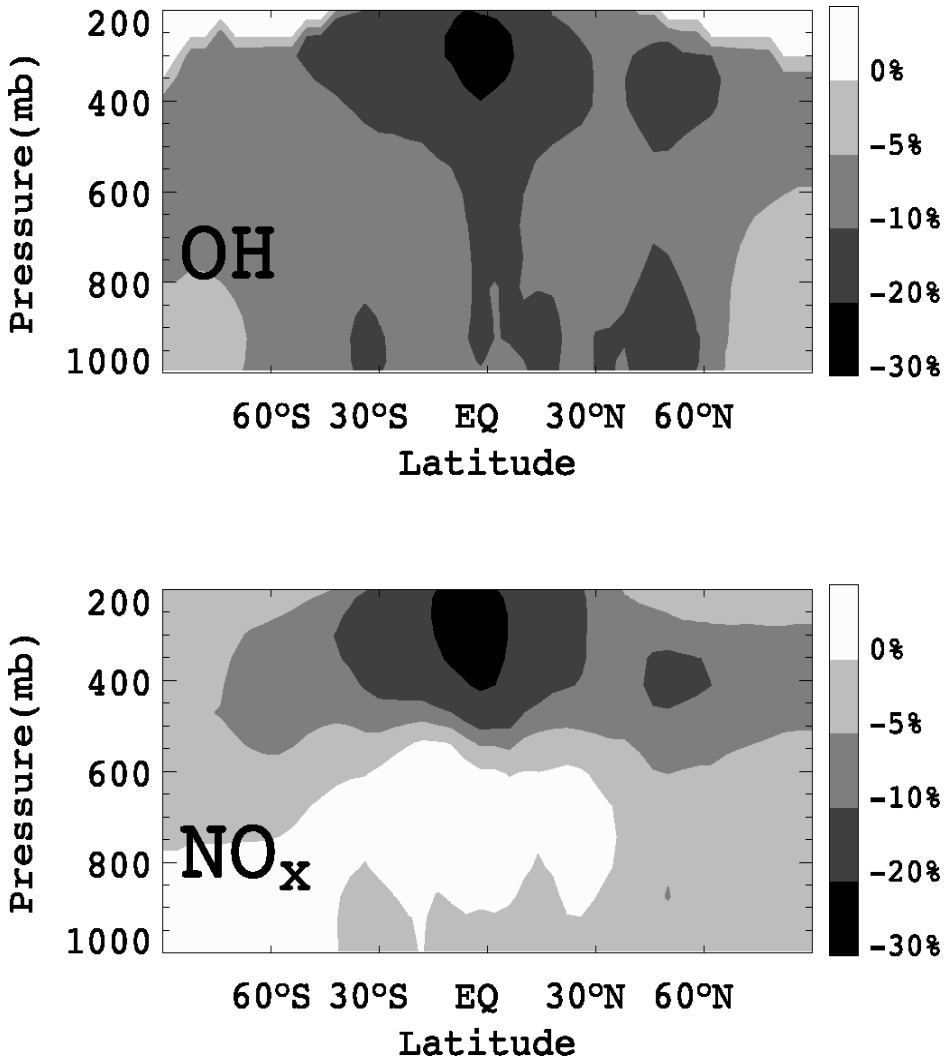


Figure 2.11. Percent changes of annual and zonal mean concentrations of OH and NO_x due to the increase of the a posteriori isoprene and the other biogenic emissions.

2.7 Conclusions and Discussion

Atmospheric VOCs play an essential role in the tropospheric chemistry. Globally isoprene accounts for a major fraction of the reactivity of VOCs. Current biogenic isoprene sources are highly uncertain due to limited global measurements, particularly over the tropics. We have presented the first Bayesian inverse modeling analysis of the global HCHO column measurements by GOME in order to evaluate the global isoprene emissions during September 1996 – August 1997. Different HCHO sources are explicitly taken into account in the inversion using GEOS-Chem as the forward model.

We selected 8 regions with high signal-to-noise ratios in GOME measurements to conduct the inversion. To facilitate inverse modeling, we applied archived OH, NO, and O_3 concentrations from the standard simulation to estimate the sensitivities of HCHO columns to biogenic emissions from 10 vegetation groups, biomass burning, and industrial VOCs. Sensitivities are close to linearity in the range of emission changes in this study. The largest deviations of ~ 8% from linearity are found over Southeast Asia and Africa.

The a priori simulation greatly underestimates global HCHO columns over the 8 regions (model bias: -14 – -46%, R: 0.52 – 0.84). The a posteriori results show generally higher isoprene and biomass burning emissions. Comparison between the a priori and the a posteriori HCHO source parameters for the 8 regions shows some general tendencies (Table 2.3). First, isoprene emissions from agricultural lands, tropical rain forest, tropical seasonal forest, and rest of the ecosystems are increased in almost all regions. Second, isoprene emissions from dry evergreen and crop/woods are reduced in most regions. Despite those tendencies, the a posteriori changes still depend on the continents. The

HCHO sources from biomass burning emissions are increased in all regions. The biomass burning HCHO enhancements are > 400% over South America and East Asia, >200% over Southeast Asia, and ~50% over Africa. The a posteriori simulation improves the model bias for all regions (model bias: -3.6 – -25%, R = 0.56 – 0.84).

There is a significant discrepancy between the seasonality of GOME measured and GEOS-Chem simulated HCHO columns over the northern equatorial Africa. We attribute this problem to the incorrect seasonal cycle in surface temperature used in GEOS-Chem. As a result, isoprene emissions over the region are overestimated. We also find that the model cannot reproduce the observed seasonal HCHO column increase from winter to summer over Southeast Asia and India. A major limitation of this study is due to the large uncertainties of the GOME HCHO column measurements resulting in relatively large uncertainties in the a posteriori emission estimates.

The a posteriori estimate of the annual global isoprene emissions of 566 Tg C yr^{-1} is about 50% larger than the a priori estimate. Table 2.5 shows the ratios of the a posteriori base emission rates to those of GEIA for different vegetation types and continents (the monthly mean a posteriori isoprene emissions are shown in Figure 2.12 and corresponding data are also available at http://apollo.eas.gatech.edu/data/isoprene_05).

The increase of global isoprene emissions significantly perturbs tropospheric chemistry, decreasing the global mean OH concentration by 10.8% from 1.06 to 0.95×10^6 molecules/cm^3 and increasing the tropospheric O_3 burden by 1.5% from 328 to 333 Tg. The atmospheric lifetime of CH_3CCl_3 increases from 5.2 to 5.7 years in closer agreement with the estimate by Prinn *et al*. (2001).

We find that the a posteriori global isoprene annual emissions are generally higher at mid latitudes and lower in the tropics when compared to the GEIA inventory (Guenther *et al*., 1995). The large reduction (a factor of 3) of isoprene emissions for some tropical ecosystems based on limited in situ measurements as used in the a priori simulations appears to be supported only for tropical rain forest in South America and tropical seasonal forest in Africa (Table 2.5). Our results indicate large variations in the reduction factor ranging from 0 to 250% depending on region and ecosystem

In summary, the a posteriori results suggest higher isoprene emissions than a priori for agricultural lands, tropical rain forest, tropical seasonal forest, and rest of the ecosystems and lower isoprene base emissions for dry evergreen and crop/woods. The a posteriori biomass burning HCHO sources increase by a factor of 2 – 4 in most regions with significant emissions except for India (only ~10%). The industrial HCHO sources are higher by ~20% except for East Asia and India (~60%). Lastly, the a posteriori uncertainties of emissions, although greatly reduced, are still high (~90%) reflecting the relatively large uncertainties in GOME retrievals.

We did not include the kinetics uncertainty of HCHO yields from isoprene oxidation in the inversion. Further studies are merited on how to properly account for this uncertainty. Given the complexity of biogenic emissions and the enormous biodiversity in ecosystems, improved in situ measurements are likely to be available only for specific regions like North America and Europe, where the research capability and resources are up to this difficult task. On the global scale, however, more accurate HCHO or other proxy observations from the next generation satellites are necessary to improve the biogenic emission inventories.

Recently, with help of advancing observational skills and computational capabilities, a comprehensive study to address both the updating estimates of global isoprene emissions with improved resolutions based on space-based measurements is possible. For instance, there are HCHO columns available with much better resolution, including the Ozone Monitoring Instrument (OMI), launched in 2004 aboard the NASA Aura satellite, providing 13 x 24 km^2 resolution as well as daily global coverage. The retrieval of those columns would allow advanced study possible with a greater time and spatial scale. The simulated shape factor with comparable resolution is necessary in order to better calculate AMF for the new column measurements. The advanced CTM (e.g, GEOS-Chem) could be employed. Recent version of GEOS-Chem uses updated assimilated meteorological variables from NASA Goddard Earth Observing System (GEOS-4) and has 1˚ x 1˚ simulation capability with 48 sigma vertical layers. Additionally aerosol simulation schemes can consider the aerosol extinctions despite some discrepancies with measured ones (Heald *et al*, 2005; Millet *et al*., 2005. The new version of CTM should be applied for retrieval of HCHO vertical columns from the new generation of satellite instruments for the years when both observations and simulation are available.

There is also improvement in a priori bottom-up isoprene emission estimates (estimation from small scale emission factors) by Guenther *et al*. (2005) using Model of Emissions of Gases and Aerosols from Nature (MEGAN), which considered more sophisticated emissions schemes. Specifically, Guenther considered the emissions with a diverse time scale factor, up to hourly emission activities. Coupled with a more realistic canopy radiation model including a leaf energy balance, the adjusted average total global isoprene emissions were similar to his previous estimates (Guenther *et al*., 1995), but

with changes in emission distribution, the reported estimates lead to a significant improvement in the uncertainty. Combined with those improved measurements and priori information and with more sophisticated and flexible inversion models such as nonlinear Bayesian inverse methods, the accuracy of emission inventories could be improved and that leads to better understanding of complicated interactions between biosphere and atmosphere.

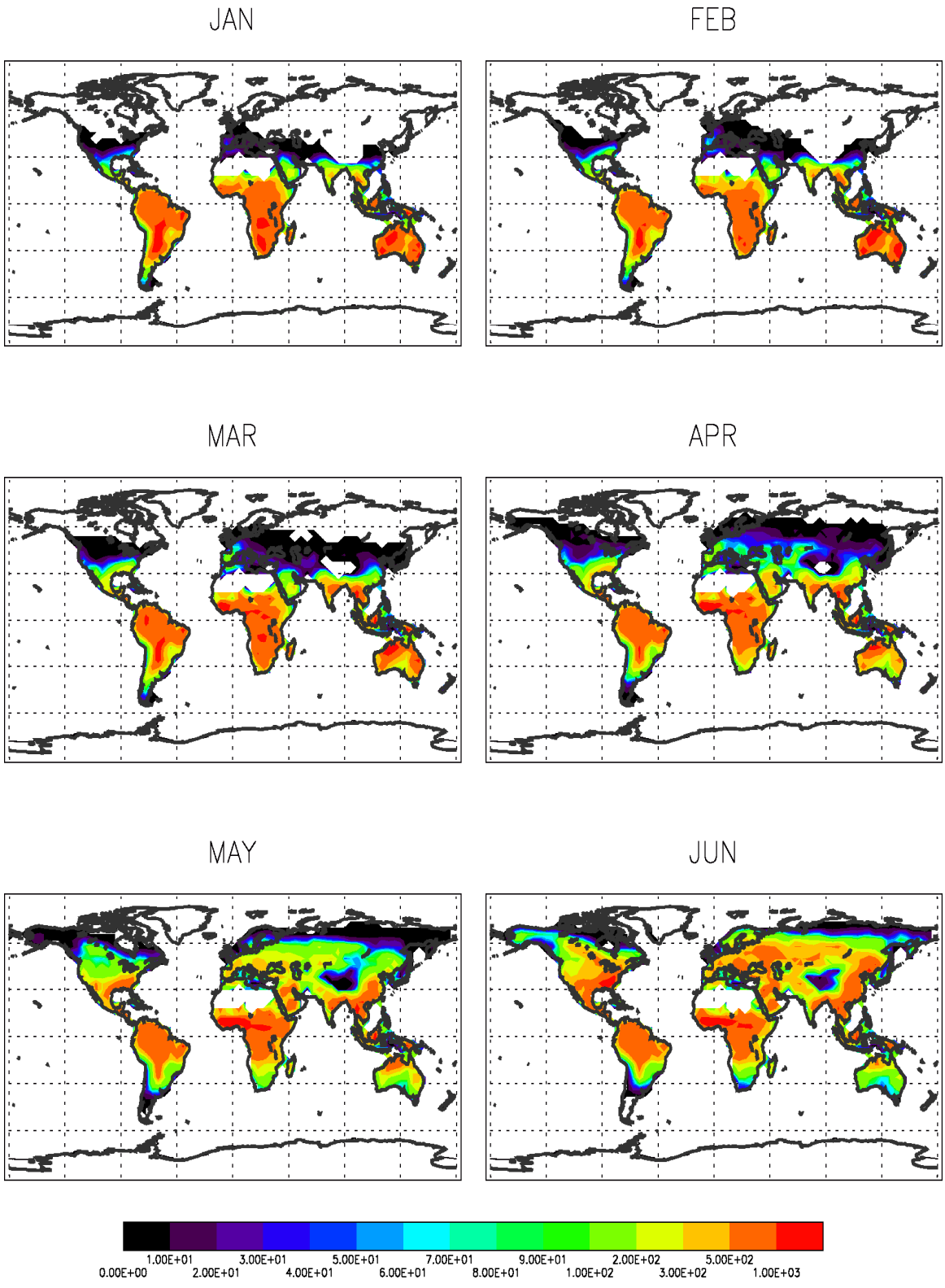


Figure 2.12. Monthly mean a posteriori global isoprene emissions (mgC/m^2 month) with 4° x 5° resolution.

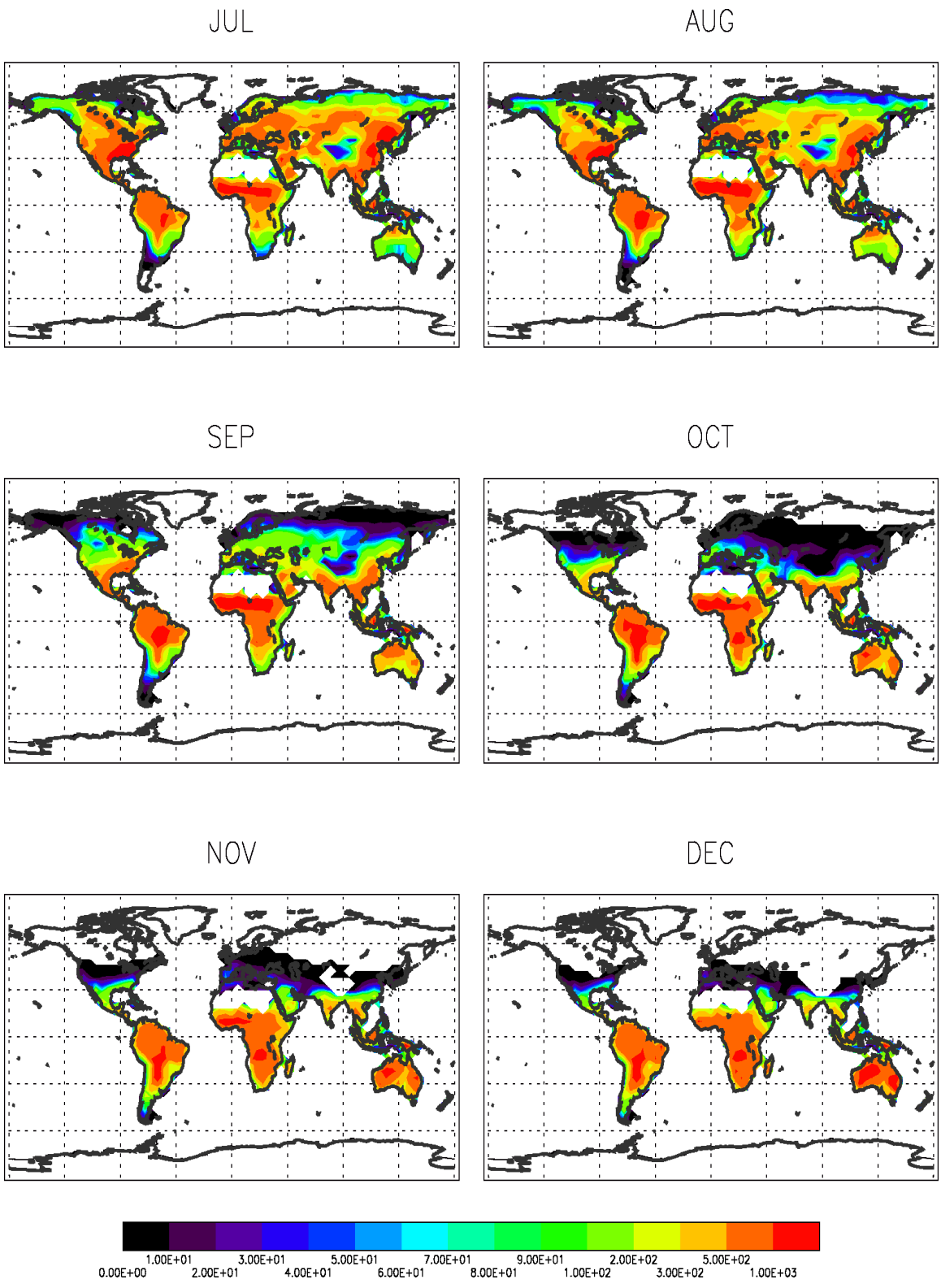


Figure 2.12. (Continued) Monthly mean a posteriori global isoprene emissions (mgC/m^2 month) with 4° x 5° resolution.

CHAPTER III

SOURCE CHARACTERISTICS OF OXYGENATIED VOLATILE ORGANIC COMPOUNDS AND HYDROGEN CYANIDE[1]

3.1 Introduction

Oxygenated volatile organic compounds (OVOCs) are a critical component of the photochemical system of the troposphere (e.g., Singh *et al.*, 2004). Previous studies found significant emissions of OVOCs from the biosphere, which affect not only trace gas composition and chemistry but also aerosols and subsequently the radiative budget of the atmosphere (e.g., Fuentes *et al.,* 2000; Guenther *et al.*, 2000; Fall, 2003). There have been many efforts to estimate the global budgets of OVOCs (Guenther *et al.,* 1995; Galbally and Kristine, 2002; Heikes *et al.,* 2002; Jacob *et al.,* 2002; Singh *et al.*, 2004). Despite these efforts, there are still large uncertainties and many unknowns.

Hydrogen cyanide (HCN) is an often used indicator for biomass burning (e.g., Rinsland *et al*., 1998). Additionally, biogenic processes are known to produce HCN and CH_3COCH_3 simultaneously (Conn *et al.,* 1991). A specific metabolic process is cyanogenesis in plants (e.g. food crops, clovers, and eucalyptus leaves), which is part of plant self-defensive activity to fend off herbivores during growing seasons (Fall *et al.,* 2003; Gleadow and Woodrow, 2000). Global estimate of the biogenic HCN production is highly uncertain partly because this complicated biochemical process is not well understood over the diverse terrestrial ecosystems.

---

[1]This chapter is extension of "Source characteristics of oxygenated volatile compounds and hydrogen cyanide," submitted to *Journal of Geophysical Research* in May 2006. Authors are **Changsub Shim**, Yuhang Wang, Hanwant B. Singh, Donald R. Blake, and Alex B. Guenther.

Furthermore, direct HCN measurements alone cannot be used to identify biogenic sources because of its long lifetime. In order to identify and quantify major source contributions to OVOCs and cyanides, we analyze aircraft measurements made during the TRACE-P, PEM-Tropics B, and INTEX-NA experiments (Raper *et al.,* 2001; *Jacob et al.*, 2003; Singh *et al*., 2002). We apply factor analysis using the positive matrix factorization (PMF) method to analyze these measurements. We show the characteristics of each source factor and calculate their contributions to the OVOCs and cyanides. Finally, we compare these results with the global budgets estimated by the previous studies.

3.2. Methodology

We choose a suite of chemicals, including methanol (CH_3OH), acetone (CH_3COCH_3), acetaldehyde (CH_3CHO), ethyne (C_2H_2), ethane (C_2H_6), i-pentane (i-C_5H_{12}), carbon monoxide (CO), methyl chloride (CH_3Cl), and bromoform ($CHBr_3$). The suite of chemicals is chosen to resolve the sources for OVOCs and cyanides; the limited tracer set is not meant to resolve different industrial/urban emission sources. Methanol has large terrestrial biogenic emissions (Fall and Benson, 1996; Galbally and Kristine, 2002). Major CH_3COCH_3 sources are from the terrestrial biosphere (Pouton, 1990; Jacob *et al.*, 2002; Singh *et al*., 2004), biomass burning (Jost *et al.*, 2003), and the ocean (Jacob *et al.*, 2002). CH_3CHO has known biogenic and oceanic sources (Harry and Kimmerer, 1991; Singh *et al.,* 2004), although the uncertainty in the budget is large. C_2H_2 and CO are good tracers for combustion. C_2H_6 is a good tracer for the usage of fossil fuel. i-C_5H_{12} is a unique tracer for gasoline evaporation and gasoline powered motor vehicle emissions

and hence the emissions from urban regions in general. i-C_5H_{12} is not included in the PEM-Tropics B analysis because its concentrations in the remote tropical Pacific are generally below the detection limit. CH_3Cl is a typical tracer for biomass burning and biogenic emissions (Yoshida *et al.,* 2004 and references therein). $CHBr_3$ is a tracer for ocean emissions (Singh *et al.,* 2004). In addition, HCN and CH_3CN, available in TRACE-P and INTEX-NA, are included. They are generally considered biomass burning tracers (de Gouw *et al*., 2003; Li *et al.*, 2003; Singh *et al.*, 2003a).

To investigate the major source factors contributing to the variability of these tracers, the method we use is PMF, which has been used previously to identify source factors of aerosols and trace gases (Paatero, 1997; Wang *et al.,* 2003; Liu *et al.,* 2005). The factors are obtained from examining the covariance structure of the datasets (Paatero, 1997). In the analysis, we selected only coincident measurements of the chemicals. Missing data are not used in order to reduce the uncertainty in the analysis.

The data matrix *X* of *m* measurements by *n* tracers can be decomposed by PMF for *p* factors as

$$X = GF + E \tag{3.1}$$

or

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij} \tag{3.2}$$

$$i = 1, \ldots, m; j = 1, \ldots, n; k = 1, \ldots, p.$$

where the *m* by *p* matrix G is the mass contributions of k^{th} factor to i^{th} sample (factor score), the *p* by *n* matrix F is the gravimetric average contributions of k^{th} factor to j^{th}

chemical species (factor loadings), and the *m* by *n* matrix E is the error. We also use the explained variation (EV),

$$EV_{kj} = \sum_{i=1}^{m} | g_{ik} f_{kj} | / \left[\sum_{i=1}^{m} \left(\sum_{k=1}^{p} | g_{ik} f_{kj} | + | e_{ij} | \right) \right] \quad (3.3)$$

to define the relative contributions of each factor to chemical species since the mixing ratios of the compounds are directly comparable. The analysis approach is the same as used by Wang *et al.* (2003b).

As commonly found in principal component and factor analyses, the minimization of the error matrix does not land to a unique decomposition of matrix X to matrices G and F. Matrix rotation is one way to explore the constraints of the covariance structure. As in the work by Wang *et al.* (2003b), the results presented here are not sensitive to rotation. For factor analysis, a key parameter is the choice of factor number. By evaluating the error matrix E (Paatero, 1997; Lee *et al*., 1999; Paatero *et al.*, 2002), a range of factor numbers is determined. We then inspect the PMF results and determine the number of factors (with necessary rotation) that leads to clear separation of tracer signals. Following this procedure, we determined 6 factors for TRACE-P and 5 factors for PEM-Tropics B, which provide most physically meaningful results.

In the work by Wang *et al*. (2003b), PMF analysis was used to identify air mass characteristics. That is the first objective of this work. The second objective is to investigate if the tracer signals in factor contributions can be used to gauge our understanding of emissions of these trace gases. In this case, the non-negativity in factor contributions calculated using an alternating least-square method in PMF (Paatero, 1997)

offers a major advantage over less advanced techniques because source contributions by definition can only be positive. In order to attempt the second objective, the inversion of sources using factor analysis makes the assumption that factor profiles represent a linear combination of source contributions. The assumption is obviously false for chemicals with large secondary sources. It is also false if the chemicals have short lifetimes in the troposphere. The influence of transport is difficult to quantify in factor analysis.

We note here two specific problems arising from transport. First, the anti-correlation driven by source location difference (for example, the often observed anticorrelation between ozone transported from the stratosphere and CO emitted from the surface), which tends to manifests in principal component analysis, does not pose a problem in PMF analysis. Second, transport processes do not "create" air masses with clean tracer characteristics; mixing during transport reduces source signals instead. As a result, the source characters cannot always be resolved by factor analysis. In this work, we show that the tracer source characteristics can be resolved for TRACE-P and PEM-Tropics B. However, transport/mixing processes during INTEX-NA results in mixed source characteristics. PMF cannot clearly resolve the source contributions in that dataset.

Despite all the assumptions, we will attempt to compare source contributions inferred from PMF analysis of the two datasets with current source estimates. The empirical approach is advantageous over 3-D model studies when the source types, distributions, and strengths for the chemicals of interests are not well understood. Furthermore the PMF results are not affected by uncertainties in 3-D model chemistry and transport. Figure 3.1 shows the measurement regions during TRACE-P, PEM-Tropics B, and INTEX-NA experiments.

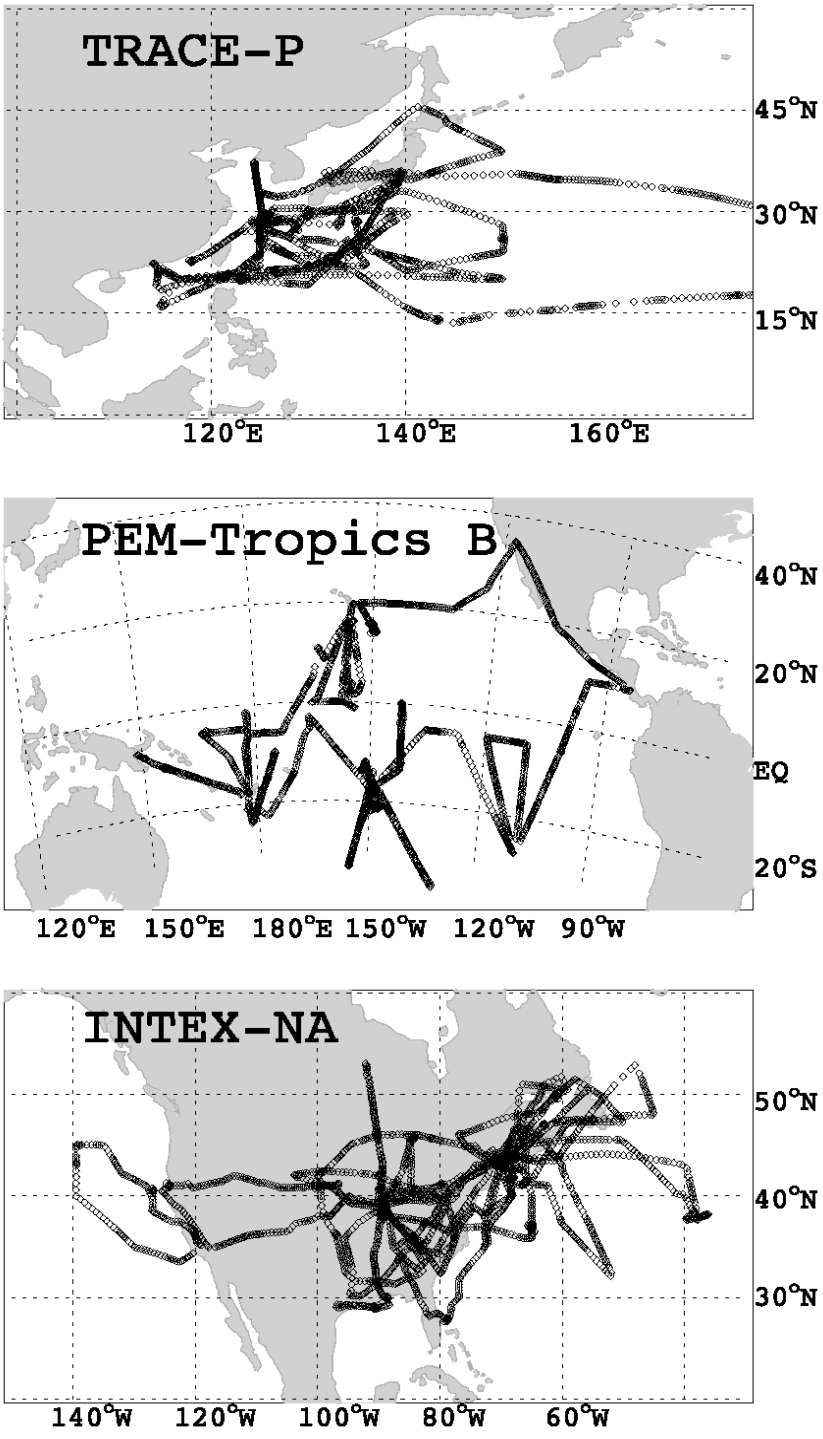


Figure 3.1. Location of aircraft measurements during TRACE-P (top), PEM-Tropics B (middle), and INTEX-NA (bottom).

3. 3. Results

3.3.1 TRACE-P

The TRACE-P experiment was conducted to investigate the Asian outflow to the Pacific during springtime (March – April, 2001) (Jacob *et al.,* 2003). There are 386 coincidental measurements over the region of 15°–45°N and 114°E–124°W. 50% of the selected data is distributed below 3km. Figure 3.2 shows the EV profiles of 6 factors, and Figure 3.3 shows the absolute variability explained by each factor.

We label each factor based on its key chemical signatures. The "terrestrial biogenic" factor is characterized with large signals in CH_3OH (80%, 862 pptv) and CH_3COCH_3 (40%, 215 pptv). Terrestrial biosphere is the largest known source for CH_3OH (e.g., Heikes *et al.,* 2002) and a substantial source for CH_3COCH_3 (e.g., Jacob *et al.*, 2002). HCN is often thought to originate from biomass burning. However, the terrestrial biogenic factor also explains a large fraction of the HCN variability (41%, 72 pptv). Cyanogenesis is likely to be a major biogenic process emitting HCN and its by-product, CH_3COCH_3 (40%, 215 pptv) from a variety of plants (Pouton, 1990; Vetter, 2000; Fall, 2003). The biogenic factor explains 30% of CH_3CN (27 pptv). However, this fraction is not so significant considering its large measurement uncertainty of 35 pptv (Table 3.1). The biogenic CH_3CHO variability (35%, 79 pptv) is much greater than its measurement uncertainty (Table 3.1), providing support of its biogenic production (Harry and Kimmerer, 1991; de Gouw *et al.*, 1999; Karl *et al.,* 2002).

The "biomass burning" factor is loaded with largest signals of CH_3CN (69%, 63 pptv) and CH_3Cl (61%, 40pptv), two chemicals with large biomass burning sources (e.g., Singh *et al*., 2003a; Yoshida *et al.*, 2004). HCN, known to have a large biomass burning source

(e.g., Li *et al.,* 2000), also has a large signal (30%, 53 pptv). The signals of CH_3COCH_3 (19%, 104 pptv) and CO (23%, 22 ppbv) are also significant in this factor.

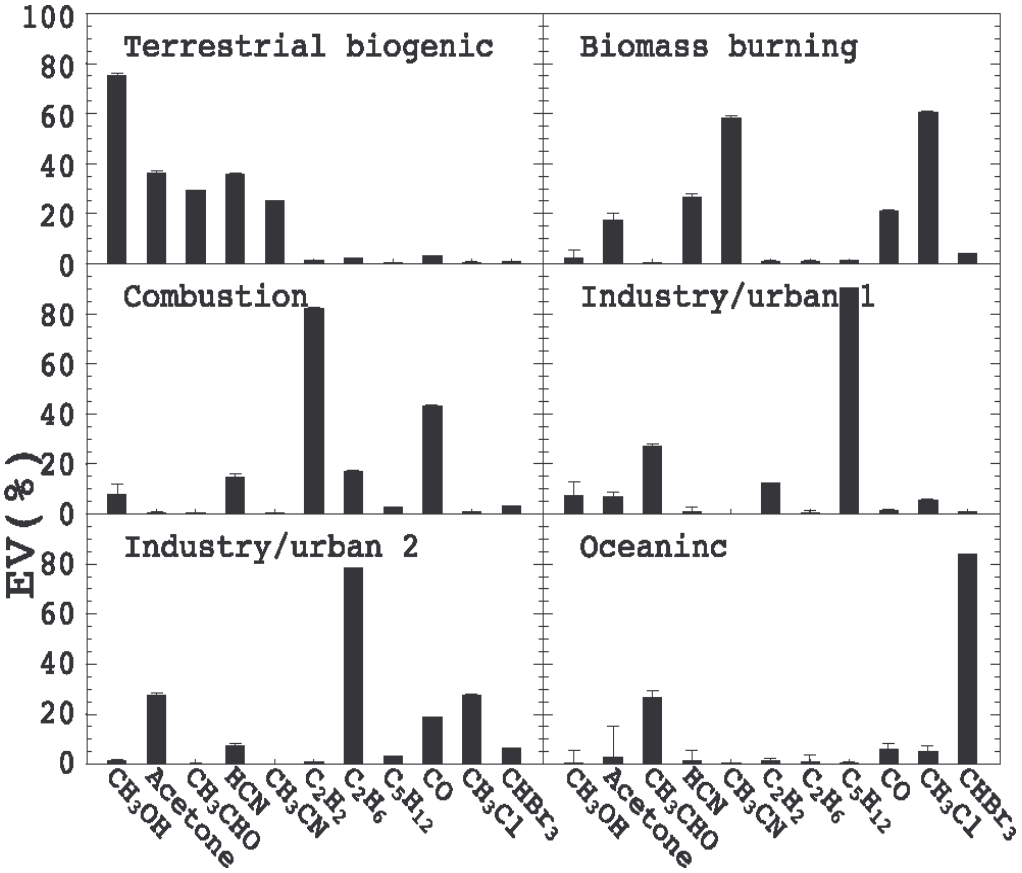


Figure 3.2. Explained variations of chemical compounds in the 6 factors for TRACE-P. The factors are named based on the chemical characteristics.

Two key combustion tracers, C_2H_2 (82%, 392 pptv) and CO (46%, 44 ppbv), appear in the "combustion" factor. Both gases are emitted from biomass burning and motor vehicle combustion. The similar correlations appear in the covariance structure of the dataset, resulting in one factor that represents a mixture of two sources. The source ambiguity, however, does not affect the gases of interest in this study.

Table 3.1. Relative and average absolute measurement uncertainties.

	CH_3OH	CH_3COCH_3	CH_3CHO	HCN	CH_3CN	C_2H_2	C_2H_6	C_5H_{12}	CO	CH_3Cl	$CHBr_3$
Relative uncertainty	10%	10%	10%	~25%	~25%	1%	1%	1%	7%	2%	5%
TRACE-P[1]	110	90	30	50	35	5	15	2	7	12	0.03
PEM-Tropics B[1]	90	40	10	-	-	3	3	-	4	12	0.03

[1]Unit is pptv except for CO (ppbv).

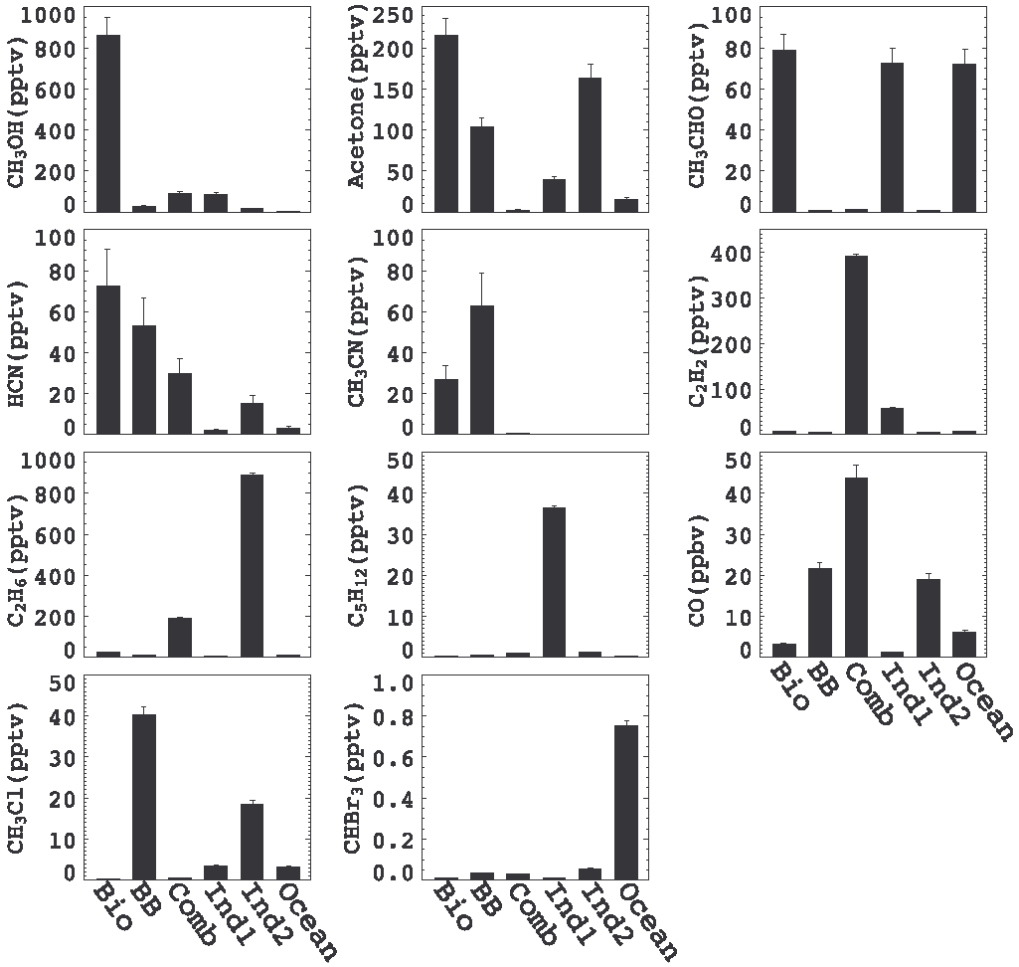


Figure 3.3. **The variabilities in mixing ratio explained by each factor for TRACE-P. "Bio" denotes terrestrial biogenic; "BB" denotes biomass burning; "Comb" denotes combustion; and "Ind" denotes industry/urban. The vertical bars show the uncertainties of the PMF analysis.**

The industry and urban pollution contributions are split into two factors. The first factor has significant signals of CH_3CHO (32%, 73 pptv), C_2H_2 (12%, 58 pptv), and i-C_5H_{12} (91%, 37 pptv) (industry/urban 1), while the second factor shows notable variabilities of C_2H_6 (78%, 888 pptv), CH_3COCH_3 (30%, 163 pptv), and CO (20%, 19 ppbv) (industry/urban 2). The first factor is more characteristic of urban emissions while the second is more of industrial emissions. The CH_3CHO variability explained by the industry/urban 1 factor is much higher than the primary biomass burning (5%) or industrial (< 1%) CH_3CHO source estimated by Singh *et al.* (2004), which likely reflects secondary production of CH_3CHO from VOC oxidation. The variability of CH_3Cl (28%, 18 pptv) in the industry/urban 2 factor is much higher than the global industrial source fraction of (4%) (McCulloch *et al.,* 1999), reflecting the influence by biofuel combustion in relatively fresh industrial plumes sampled during TRACE-P. The most prominent component in the "oceanic" factor is $CHBr_3$ (84%, 0.75 pptv). This factor also explains 32% of CH_3CHO (72 pptv), supporting the notion of substantial CH_3CHO oceanic emissions (Singh *et al.,* 2003b).

Previous studies suggested large sources of CH_3Cl from the terrestrial biosphere (Yoshida *et al.*, 2004 and references therein). However, the CH_3Cl variability explained by the terrestrial biogenic factor is negligible. The fairly small variability of the biogenic CH_3Cl likely reflects its dispersed terrestrial sources in the tropics (Yoshida *et al.,* 2004) and mixing during long-range transport. Considering the small contribution to CH_3Cl by its largest known source from the tropical terrestrial ecosystems to this factor, the much larger HCN signal here in turn likely indicates that measured enhancements originate from the forests at mid latitudes, not in the tropics. Figure 3.4 shows the kinematic 5-day

back-trajectories of biogenic and biomass burning factors calculated with the European Centre for Medium-Range Weather Forecasts (ECMWF) meteorological fields (Fuelberg *et al*, 2003). The biomass burning factor is associated with transport from tropical regions, while the terrestrial biogenic factor is related to transport from dispersed continental locations (Figure 3.4).

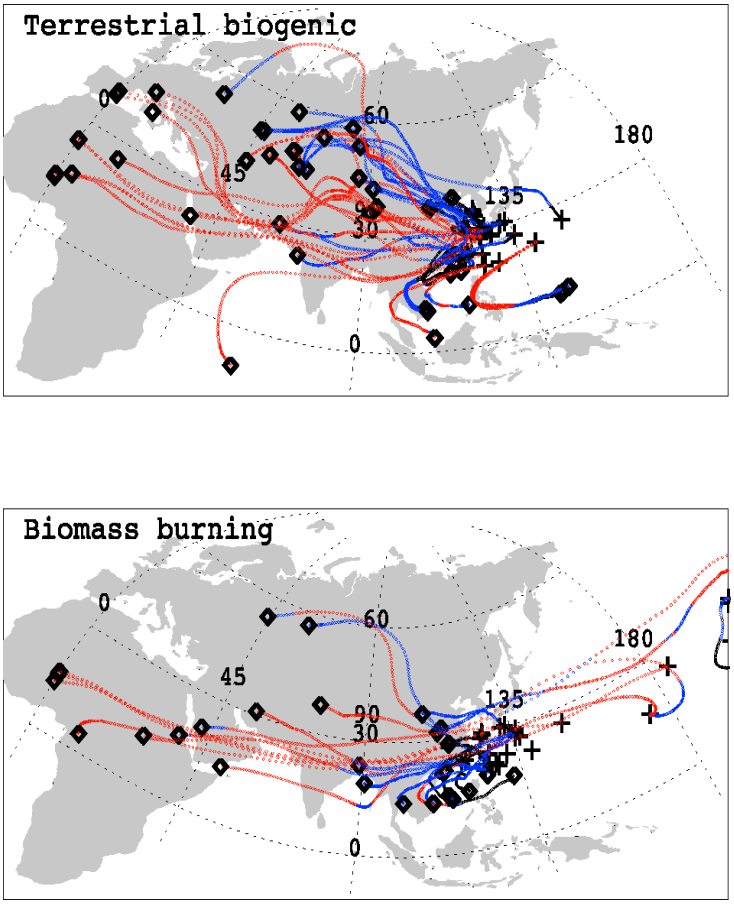


Figure 3.4. Five day back trajectories of data points that rank in the top 10% by factor score for the terrestrial biogenic (top) and biomass burning (bottom) factors in TRACE-P. The crosses show the locations of measurements. The back trajectory locations are represented by dots with colors indicating heights (black (< 2 km), blue (2 – 6 km), and red (> 6 km)) and the end point is marked by a diamond.

To further investigate the characteristics of the biomass burning factor, we examine how the factor correlates with latitude, altitude, and C_2H_6/C_3H_8 ratio. Higher C_2H_6/C_3H_8 ratios indicate more aged air masses (Wang and Zeng, 2004). The correlation is strongest with C_2H_6/C_3H_8 ratio (r = 0.73), indicating aged biomass burning air masses. The positive correlation with altitude (r = 0.59) and negative correlation with latitude (r = -0.31) indicate transport of aged biomass burning air masses from tropical regions at high altitude after initial plume rise driven by combustion heating, consistent with backtrajectories (Figure 3. 4).

The significant positive correlations of biomass burning factor with altitude and C_2H_6/C_3H_8 ratio are unique (Table 3.2; Figure 3.5). The terrestrial biogenic factor has no apparent correlation with latitude or altitude, implying a geographically dispersed distant source. The lack of correlation with C_2H_6/C_3H_8 ratio reflects in part that the biogenic sources are not collocated with alkane sources. The two industry/urban factors show significant anitcorrelations with altitude and C_2H_6/C_3H_8 ratio, representing relatively fresh plumes from low-altitude sources. Between them, industry/urban 2 factor with more chemical signatures of industrial than urban emissions have stronger anticorrlations with altitude and C_2H_6/C_3H_8 ratio. It also has clear latitude dependence. These characteristics are consistent with more industrial activities in the northeast of China and more dispersed urban distributions in the region. The oceanic factor has large anticorrelations with altitude and C_2H_6/C_3H_8 ratio, consistent with characteristics of fresh surface emissions. The anticorrelation with C_2H_6/C_3H_8 ratio reflects in part a positive correlation of the hydrocarbon ratio with altitudes.

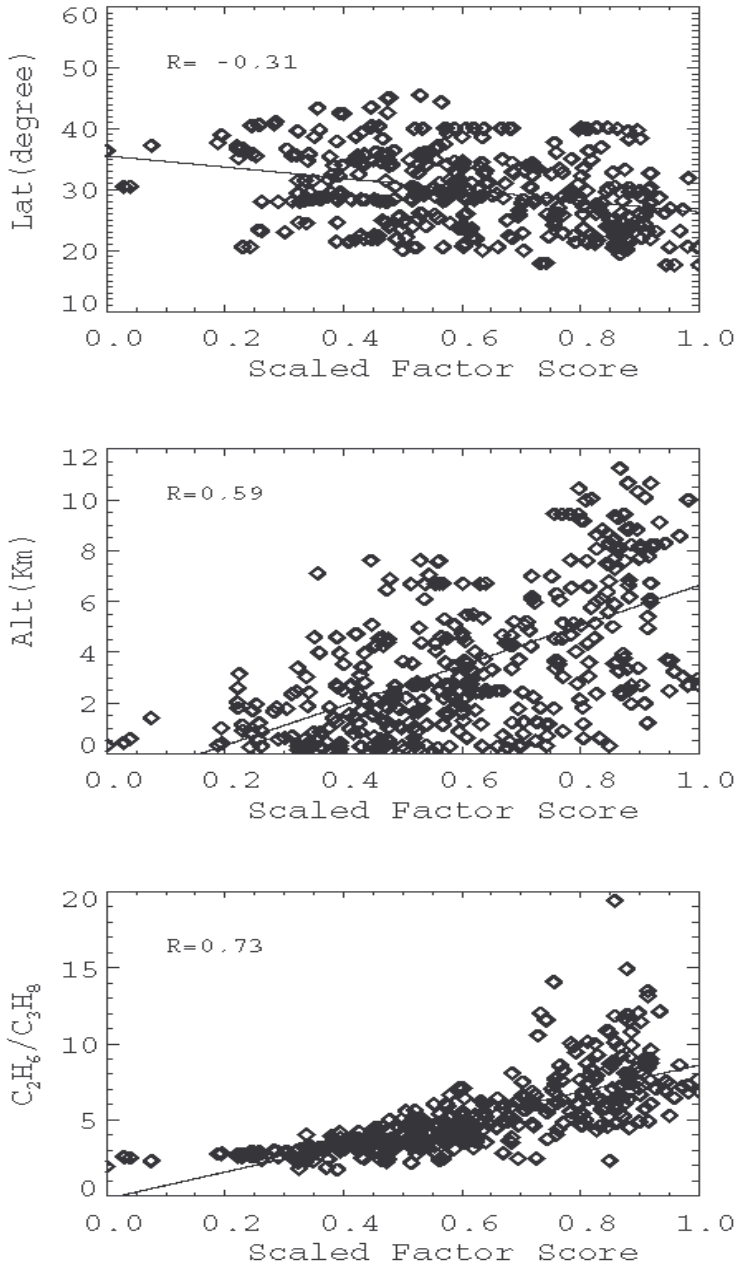


Figure 3.5. Factor correlation with latitude (upper), altitude (middle) and C_2H_6 / C_3H_8 ratio (bottom) for the biomass burning factor in TRACE-P.

Table 3.2. Correlations (R) between factor scores and latitudes, altitudes, and C_2H_6/C_3H_8 ratio for TRACE-P.

Factors	Latitude	Altitude	C_2H_6/C_3H_8 ratio
Terrestrial biogenic	0.06	0.01	0.08
Biomass burning	-0.31	0.59	0.73
Combustion	0.1	-0.21	-0.37
Industry/urban 1	0.04	-0.27	-0.39
Industry/urban 2	0.29	-0.53	-0.78
Oceanic	0.1	-0.57	-0.55

3.3.2. PEM-Tropics B

The PEM-Tropics B campaign was conducted during the wet season over the tropical Pacific (March – April, 1999) (Raper *et al.,* 2001). 581 coincident measurements of 8 chemicals over the region of 36°S–35°N and 107°W–162°E are analyzed. 63% of the selected data is distributed above 5 km. We find no significant bias in the spatial distributions of selected datasets compared to the entire datasets.

Five factors are resolved by PMF (section 2). Figure 3.6 shows the EV profiles and Figure 3.7 shows the absolute chemical variability. Compared to TRACE-P (Figs. 2 and 3), the tracer variability during PEM-Tropics B is much lower, leading to more ambiguity in interpretation of the analysis results because observed variabilities were at times smaller than measurement uncertainties (Table 3.1). As one could expect, the tracer variabilities of industy/urban related factors in PEM-Tropics B are much smaller than those in TRACE-P, which reflects the sampling of photochemically aged air plumes

during PEM-Tropics B.

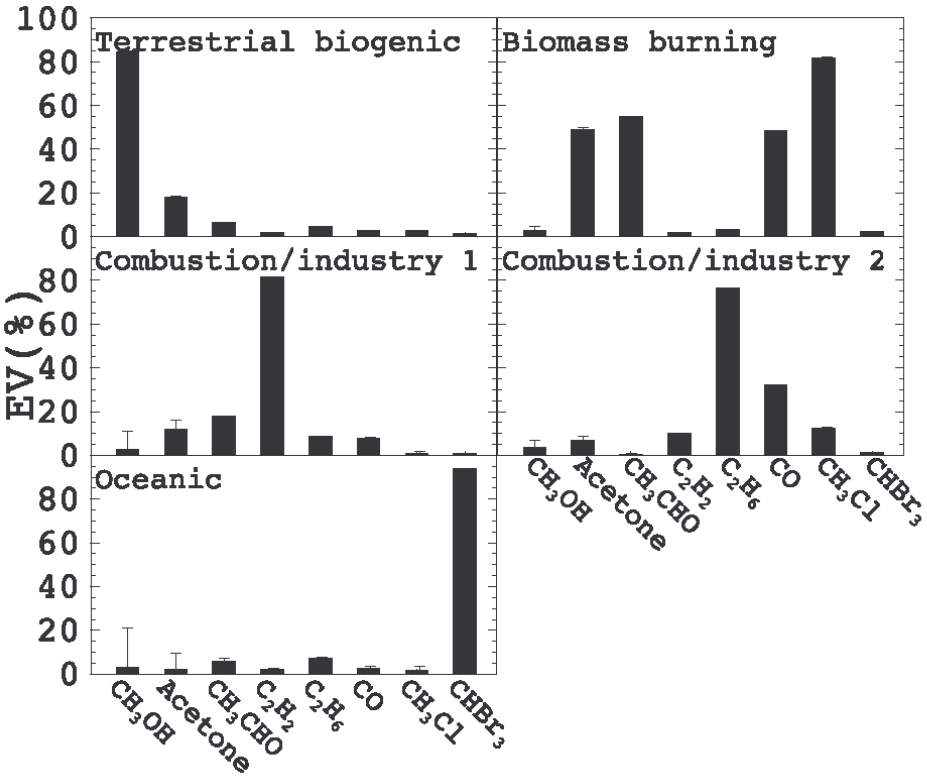
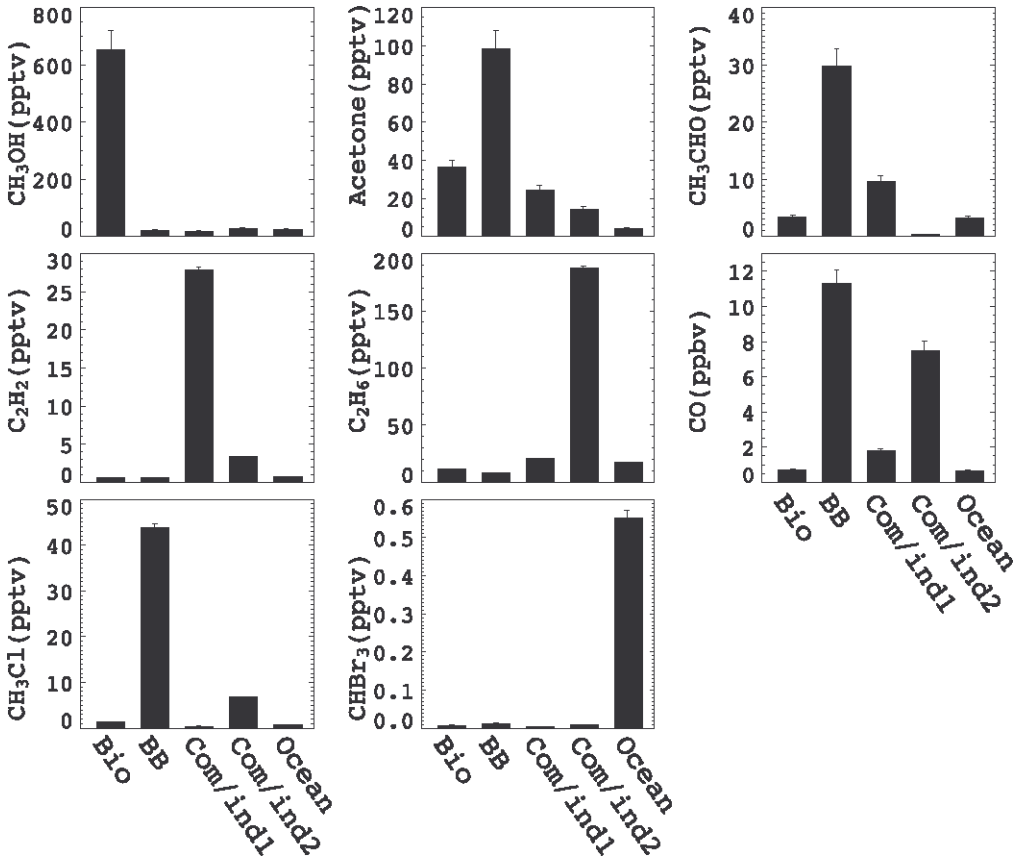


Figure 3. 6. Same as Figure 3.1 but for PEM-Tropics B.

As previously, we label each factor based on its key chemical signatures. The "terrestrial biogenic" factor is characterized by the large variability of CH_3OH (88%, 653 pptv). The signal of CH_3COCH_3 is below the measurement uncertainty. The "biomass burning" factor has large signals of CH_3COCH_3 (55%, 98 pptv), CH_3CHO (64%, 30 pptv), CO (52%, 11 ppbv), and CH_3Cl (82%, 44 pptv). The "combustion/industry" factors reflect a mix of biomass burning, automobile combustion, industry, and urban sources. There are substantial signals of C_2H_2 (84%, 28 pptv) and C_2H_6 (9%, 21 pptv) in

the combustion/industry 1 factor, and large signals of C_2H_6 (76%, 188 pptv) and CO (34%, 8 ppbv) in the combustion/industry 2 factor. The CH_3Cl fraction (13%, 7 pptv) in the latter factor is below the measurement uncertainty. The "oceanic" factor explains 94% of $CHBr_3$ variability (0.55 pptv) and does not show significant variability in the other chemicals.



Figrure 3.7. Same as Figure 3.2 but for PEM-Tropics B. "Com/ind1" and "Com/ind2" denote combustion/industry 1 and 2, respectively.

We analyze the correlations between factors and latitude, altitude, and C_2H_6/C_3H_8 ratio to help identify the source characteristics (Table 3.3). As for TRACE-P, the terrestrial biogenic factor shows no significant correlations reflecting geographically dispersed sources. The biomass burning factor is anticorrelated with latitude (r = -0.51), implying southern hemispheric origin particularly from Africa and Australia (Figure 3.8). As in TRACE-P, it is positively correlated with altitude and C_2H_6/C_3H_8 ratio, indicating long-range transport of aged pollutants after initial rise due to combustion heating. The combustion/industry 1 and combustion/industry 2 factors show significant positive correlations with latitude and negative correlations with C_2H_6/C_3H_8 ratio, implying most long-range transport of industrial/urban/combustion emissions from the northern hemisphere (Wang *et al.*, 2001). As in Trace-P, the oceanic factor is the only one with significant anticorrelation with altitude. It also has a week anticorrelation with C_2H_6/C_3H_8 ratio, representing a mixture of relatively emissions from the ocean with transport from C_2H_6 and C_3H_8 from northern industrial regions.

Table 3.3. Correlations (R) between factor scores and latitudes, altitudes, and C_2H_6/C_3H_8 ratio for PEM-Tropics B.

Factors	Latitude	Altitude	C_2H_6/C_3H_8 ratio
Terrestrial biogenic	0.01	-0.09	-0.14
Biomass burning	-0.51	0.16	0.45
Combustion/industry 1	0.6	-0.1	-0.54
Combustion/industry 2	0.68	-0.07	-0.57
Oceanic	0.13	-0.42	-0.17

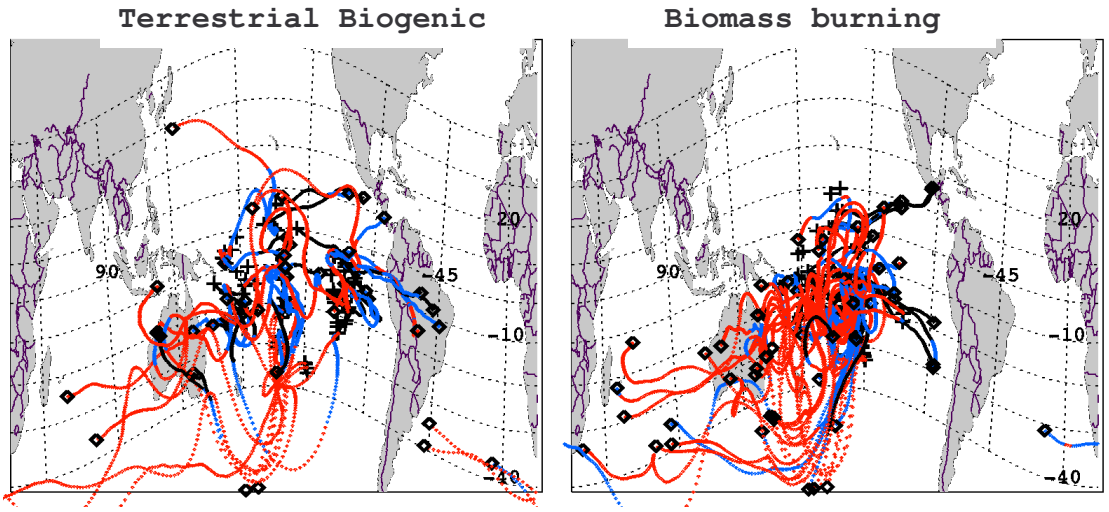


Figure 3.8. Same as Figure 3.4 but 10 day back-trajectories for PEM-Tropics B.

3.3.3. INTEX-NA

The INTEX-NA experiment is conducted to investigate the formation and transport of tracers and aerosols and inflow and outflow of pollutants over North America during the summer (July – August, 2004) (Singh *et al*., 2002). Two characteristics separate INTEX-NA from TRACE-P and PEM-Tropics B. First, mixing processes are much more active over the continent in summer than spring (TRACE-P) or over the tropical Pacific (PEM-Tropics B). Second, North American urban and industrial sources are more dispersed than over East Asia, where major industrial and urban emissions are densely located near coastal regions. As a result, we find that source-separation is much more ambiguous in INTEX-NA dataset than the other too. In the selected dataset, 35% is measured below 2 km; 50% is measured above 5 km.

We find five factors in the PMF analysis. Figure 3.9 shows the EV profiles. The factors are named after the tracer that has largest EV. The C_2H_2 is a combustion factor

with a much larger CH_3OH signal (> 25%), representing either a large industrial related CH_3OH source over North America or a mixture of combustion and biogenic emissions because we cannot define a clear biogenic factor. The CH_3Cl factor with large signals of biomass burning (CH_3Cl, CO, HCN, and CH_3CN) and biogenic tracers (CH_3OH, CH_3COCH_3, and HCN) likely represent a mixture of the two. The CH_3Br, C_2H_6, and i-C_5H_{12} factors are dominated by single species. Photochemical oxidation and mixing during INTEX-NA (summer) are considerably faster than TRACE-P (spring). Assuming similar emissions, we would expect lower concentrations in INTEX-NA than TRACE-P. Inspection of vertical profiles of trace gases we selected (Figure 3.10) shows that it is true for most species. However, concentrations of CH_3OH, acetone, and HCN are much higher in INTEX-NA than TRACE-P (Figure 3.10). We postulate that these enhancements are due at least in part to larger biogenic emissions in summer. However, the analysis method used in this study cannot be used to evaluate our hypothesis.

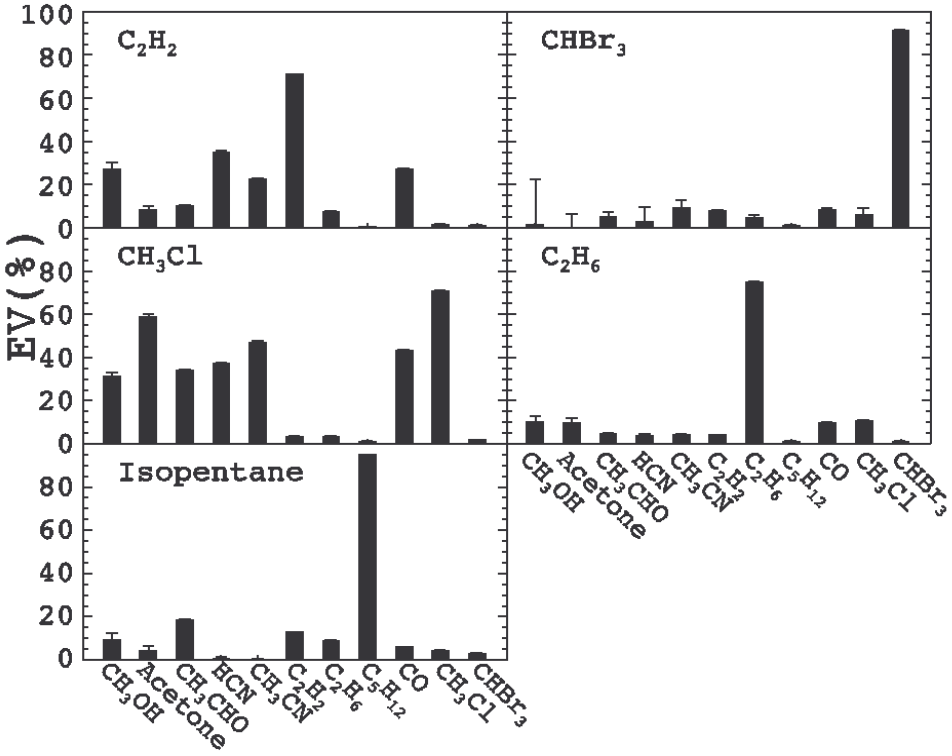


Figure 3.9. Same as Figure 3.1 but for INTEX-NA.

3.4. Comparisons to Global Estimations

Despite the limitations of PMF methods (section 2), it is still instructive to compare our results with previous global budget estimations. Additional precautions must also be heeded to prevent oversimplifications. First, our results are drawn from only two aircraft missions. Second, the factor contributions do not necessarily translate into source strengths. Third, it is generally hard to distinguish primary sources from secondary chemical production in the factor analysis. However, the following assessment points to the directions of needed research when large discrepancies are found between previous emission estimations and our results. Table 3.4 summarizes the contributions of various factors to OVOCs and HCN for TRACE-P and PEM-Tropics B; previous global estimates are also listed for comparison purposes.

The 80 – 88% variability of CH_3OH in the terrestrial biogenic factor from this study is consistent with the estimated global biogenic fractions of 62 – 81% (Heikes *et al.,* 2002; Singh *et al*., 2004; Jacob *et al*., 2005). For CH_3COCH_3, the biogenic fractions of this study (20 – 40%) are also comparable with the global estimates (27 – 53%) (Jacob *et al*., 2002; Singh *et al*., 2004).

Large terrestrial biogenic HCN variability (41%) is found during TRACE-P; it is considerably higher than the global biogenic fraction (~18%) estimated by Li *et al.* (2003, Table 3.4). The cyanogenic process in plants is known to emit both CH_3COCH_3 and HCN. It appears to be a key contributor to HCN emissions (Fall, 2003) from geographically dispersed source regions. Contribution of biomass burning to HCN (~30%) in TRACE-P is significantly smaller than the global estimates (> 80%). The industry/urban/ocean factor contributions are uncertain since they are les than the measurement uncertainty of HCN (~25%). The CH_3COCH_3 variabilities explained by industry/urban and biomass burning factors (8 – 30% and 19 – 55%, respectively) are much higher than the estimated primary sources (1 – 4% and 4 – 10%, respectively).

The total difference between PMF factor contributions and previously estimated primary sources, however, are comparable to the 30 – 50% contribution from VOC oxidation (Jacob *et al*., 2002; Singh *et al*., 2004), indicating secondary productions in those factor. There is no clear evidence for significant oceanic CH_3COCH_3 emissions, which is consistent with the recent measurements by Marandino *et al.* (2005).

The terrestrial biogenic CH_3CHO variability is quite different between TRACE-P and PEM-Tropics B (35% and 7%, respectively), likely reflecting relatively fast photochemical loss of CH_3CHO or its precursors during long-range transport to the

remote PEM-Tropics B region. The corresponding global fraction estimated by Singh *et al*. (2004) is 21%. We find large contributions from industry/urban 1 (32%) for TRACE-P and from biomass burning (64%) for PEM-Tropics B. Given the relatively short lifetime of CH_3CHO (~ 1 day), our results indicate that secondary production from VOCs, likely produced in combustion processes, is a major source of observed CH_3CHO. The oceanic CH_3CHO variability is significant during TRACE-P (32%) but not PEM-Tropics B, indicating the possibility of regional variations in oceanic CH_3CHO emissions. In comparison, the estimated global oceanic contribution is much larger at 63% (Singh *et al.*, 2004). In general, the interpretation of PMF results of short-lived species such as CH_3CHO is more ambiguous. More detailed and quantitative analysis will require 3-D chemical transport model simulations that explicitly treat the emissions, secondary production, oxidation, and transport of CH_3CHO.

Table 3.4. Comparison factor contributions with global estimates[1] of OVOCs (%)

	CH_3OH		CH_3COCH_3		CH_3CHO		HCN	
	Global estimates	PMF	Global Estimates	PMF	Global estimates	PMF	Global estimates	PMF
Biogenic	60 – 80	80 – 88	27 – 53	20 – 40	17.5	7 – 35	0 – 18	41
BB	5 – 9	3 – 16	4 – 10	19 – 55	5	1 – 68	80 – 98	30 – 47
Industry/urban	2 – 4	3 – 5	1 – 4	8 – 30	0.5	1 – 28	0 – 2	10
Ocean	0	0.5 – 3	0 – 28	2 – 4	62.5	6 – 36	0	2
Others[2]	15 – 33	-	30 – 50	-	15	-	-	-

[1]Global estimates are compiled from Heikes *et al*. (2002), Li *et al.* (2003), Singh *et al.* (2004), and Jacob *et al.* (2002, 2005).
[2]2nd productions in global estimates.

3.5. Conclusions

We apply an advanced factor analysis method, PMF, to analyze the source contributions to OVOCs and cyanides measured in three aircraft experiments. In order to identify contributions by emissions from the terrestrial biosphere, biomass burning, industry/urban regions, and ocean, we select a suite of trace chemicals including CH_3OH, CH_3COCH_3, CH_3CHO, C_2H_2, C_2H_6, i-C_5H_{12}, CO, CH_3Cl, and $CHBr_3$. Two cyanide compounds, HCN and CH_3CN are available and included in TRACE-P and INTEX-NA analyses. While feasible for TRACE-P and PEM-Tropics B measurements, PMF analysis cannot clearly separate source contributions in INTEX-NA measurements. More active transport/mixing processes during summer over North America and broad collocations of different types of sources in the region are likely the reasons.

We resolve 6 factors for TRACE-P and 5 factors for PEM-Tropics B. In TRACE-P, the terrestrial biogenic factor is characterized by significant variability of CH_3OH (80%, 862 pptv) and CH_3COCH_3 (40%, 215 pptv). This factor also shows large fraction of HCN variability (41%, 72 pptv). Backtrajectory analysis and a lack of correlation of this factor with altitude, latitude, or C_2H_6/C_3H_8 ratio imply widely dispersed continental sources. The significant variabilities of HCN and CH_3COCH_3 in this factor likely reflect emissions from cyanogenesis in diverse plants.

The biomass burning factor is characterized by large variabilities of CH_3CN (69%, 63 pptv), CH_3Cl (61%, 40 pptv), HCN (30%, 53 pptv), and CO (23%, 22 ppbv). It is the only factor with positive correlations with positive correlations with altitude and C_2H_6/C_3H_8 ratio, indicating long-range transport of aged biomass burning air masses at

high altitude. A negative correlation with latitude and backtrajectories indicate subtropical/tropical sources.

The combustion factor, characterized by variabilities of C_2H_2 (82%, 392 pptv) and CO (46%, 44 ppbv), reflects a mixture of biomass burning, industrial, and urban sources.. Significant negative correlations of this factor with altitude and C_2H_6/C_3H_8 ratio indicate relatively fresh surface emissions. There are two industry/urban factors. Industry/urban 1 factor, characterized by large variability of i-C_5H_{12} (91%, 37 pptv), is affected more by urban emissions. The second, characterized by a strong signal of C_2H_6 (78%, 888 pptv), is affected more by industrial emissions. As a result, the second factor has a stronger dependence on latitude likely due to high density of industries over the northeast of China. Both factors are negatively correlated with altitude and C_2H_6/C_3H_8 ratio, implying surface sources and transport. The oceanic factor is characterized by largest variability of $CHBr_3$ (84%, 0.75 pptv). It is clearly a surface source. The substantial CH_3CHO variability (32%, 72 pptv) in this factor implies a significant oceanic contribution to CH_3CHO.

Compared with TRACE-P, factor contributions in PEM-Tropics B are smaller because of transport over longer distance to the region. The terrestrial biogenic factor is similarly characterized by large variability of CH_3OH (88%, 653 pptv). The biomass burning factor have the large variabilities of CH_3Cl (82%, 44 pptv), CH_3CHO (64%, 30 pptv), CH_3COCH_3 (55%, 98 pptv), and CO (52%, 11 ppbv). The factor correlations with altitude, latitude, and C_2H_6/C_3H_8 ratio for these factors are similar to TRACE-P, indicating widely dispersed biogenic sources and long-range transport of aged biomass burning air masses at high altitude.

The two combustion/industry factors show large signals of C_2H_2 (84%, 28 pptv, combustion/industry 1), C_2H_6 (76%, 188 pptv, combustion/industry 2), and CO (34%, 8 ppbv, combustion/industry 2). Both factors are positively correlated with latitude and negatively correlated with C_2H_6/C_3H_8 ratio, representing long-range transport of industrial/urban/combustion emissions from the northern hemisphere. The oceanic factor characterized by $CHBr_3$ from surface emissions does not have a significant CH_3CHO signal.

We compare the source contributions from two aircraft field experiments to previous global estimations; identifying large discrepancies points to the directions of needed research. First, the terrestrial biogenic contribution of CH_3OH in this study is fairly consistent with the global biogenic emission estimates (80 – 88% and 62 – 81% respectively). Second, the terrestrial biogenic contribution of HCN in TRACE-P (41%) is substantially higher than the global emission estimates (0 – 18%), which suggests that the cyanogenesis in plants from widely dispersed regions is likely to be a major source of HCN (Figure 3.11.) in addition to biomass burning. Third, the biogenic contribution to CH_3COCH_3 variability is comparable with the global emission estimates (20 – 40% and 27 – 53%, respectively). However, there are much larger CH_3COCH_3 industry/urban and biomass burning contributions (8 – 30% and 19 – 55%, respectively) in this study than previous global emission (1 – 4% and 4 – 10%, respectively), reflecting the importance of secondary productions. We do not find large oceanic contributions to CH_3COCH_3. Fourth, considering its relatively short lifetime, the large contributions to CH_3CHO variability from industry/urban 1 (32%) for TRACE-P and from biomass burning (64%) for PEM-Tropics B imply that secondary production from combustion/industrial VOCs is

likely an important sources. The oceanic CH_3CHO contribution (10 – 32%) shows the regional dependence and it is lower than previous global emission estimates (63%). Based on this analysis, three specific areas emerge for future research: the global sources of CH_3COCH_3 and HCN from the cyanogenic process in plants, the VOC precursors to sustain CH_3CHO concentrations in industry/urban and biomass burning plumes, and the magnitudes and regional variability of oceanic CH_3CHO emissions.

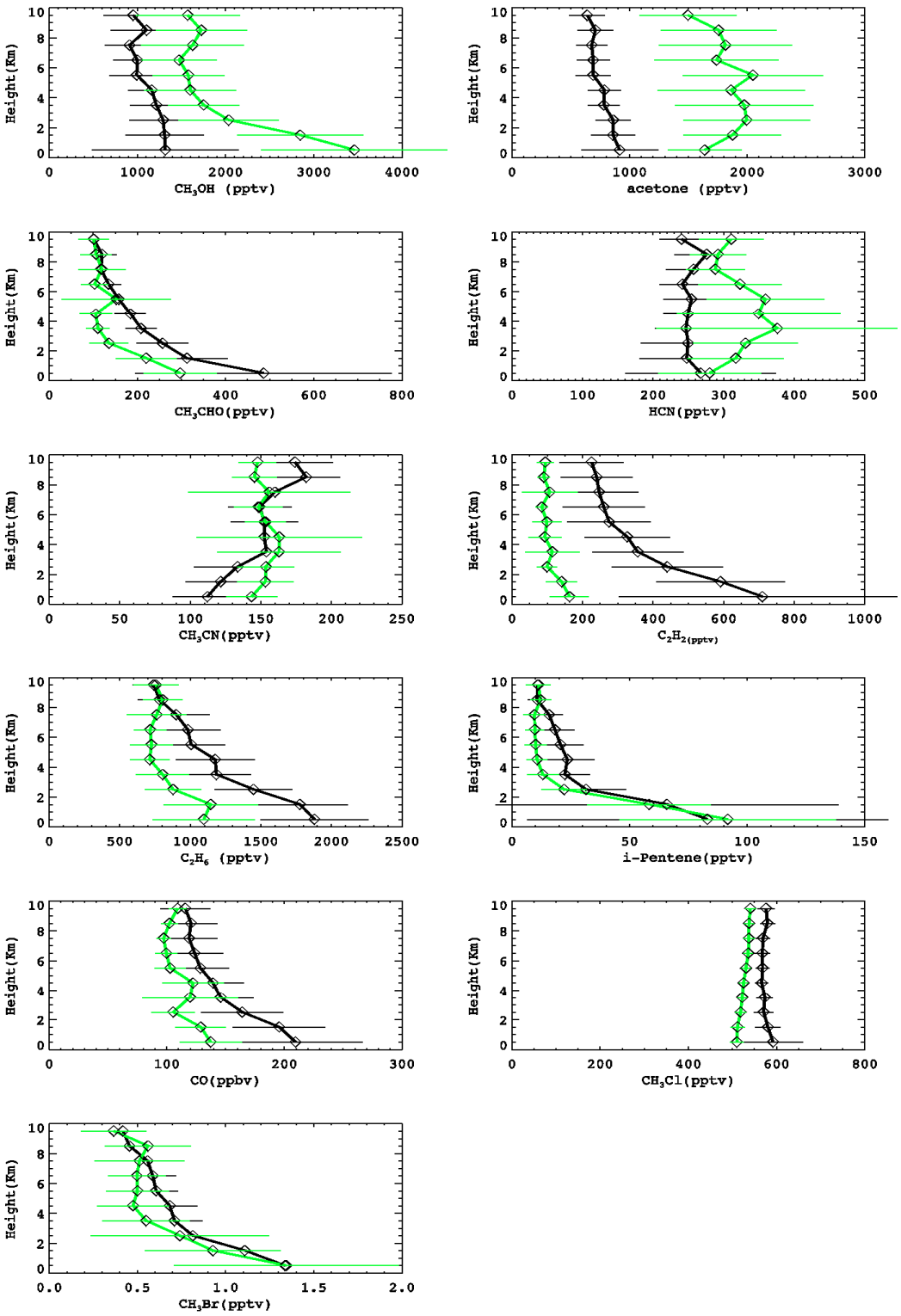


Figure 3.10. Vertical profiles of the tracers for TRACE-P (black) and for INTEX-NA (green).

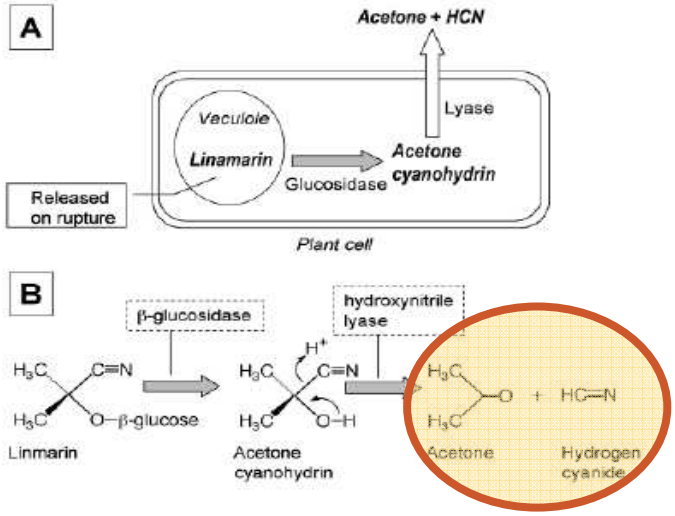


Figure 3.11. A schematic diagram of Cyanogenesis in plants (Fall *et al.*, 2003)

CHAPTER IV

INTERCONTINENTAL TRANSPORT OF POLLUTION MANIFESTED IN THE VARIABILITY AND SEASONAL TREND OF SPRINGTIME O_3 AT NORTHERN MIDDLE AND HIGH LATITUDES[1]

4.1 Introduction

Ozone in the troposphere is ether produced photochemically within the troposphere or transported from the stratosphere. Its concentrations have important environmental ramifications. At high concentrations near the surface, it is hazardous to humans and plants. It is important for maintaining the oxidation capacity of the atmosphere through its photolysis to $O(^1D)$, which reacts with H_2O to produce OH radicals. Furthermore, it is an effective greenhouse gas.

An interesting feature of tropospheric O_3 is the observed springtime increase at northern middle and high latitudes. Ozone concentrations at northern middle and high latitudes tend to peak in late spring (April and May) in the lower troposphere and in early/middle summer (May – July) in the middle troposphere (e.g., Logan, 1985). Levy *et al*. (1985) was able to simulate the springtime increase in a global three-dimensional (3-D) model, which did not incorporate tropospheric chemistry. Penkett and Brice (1986) and Liu *et al*. (1987) suggested that the increase is driven by photochemical production in the troposphere.

---

[1]This chapter is for "Intercontinental transport of pollution manifested in the variability and seasonal trend of springtime O_3 at northern middle and high latitudes," published at Journal of Geophysical Research in November 2003 (*108 (D21)* 4683, doi:10.1029/2003JD003592). Authors are Yuhang Wang, **Changsub Shim**, Nicolar Blake, Donald Blake, Yunsoo Choi, Brian Ridely, Jack Dibb, Anthony Wimmers, Jennie Moody, Frank Flocke, Andrew Weinheimer, Robert Talbot, and Elliot Atlas.

However, subsequence analysis work based on the observation correlation between O_3 and ^7Be tends to emphasize the effect of the stratospheric input (e.g. Oltmans and Levy, 1992; Dibb *et al*., 1994).

More recent analyses using photochemical models tend to emphasize the effect of tropospheric photochemical production. Photochemical box model calculations based on the measurements during the Free Tropospheric Experiment (FREETEX'96 and 98) in the Swiss Alps (at an altitude of 3.6 km) (Carpenter *et al*., 2000; Zanis *et al*., 2000) and aircraft observations during the Tropospheric Ozone Production about the Spring Equinox (TOPSE) experiment (e.g., Cantrell *et al*., 2003; Wang *et al*., 2003a) showed that tropospheric O_3 production at midlatitudes are large enough to explain the observed springtime O_3 increase at northern midlatitudes, although in situ O_3 production is ineffective at high latitudes (Wang *et al*., 2003a). Using global 3-D chemistry and transport models, Wang *et al*. (1998) showed that both factors contributed to the simulated springtime O_3 maximum at low altitudes (0 – 2 km) with one peaking in early spring and the other peaking in early summer, while Yienger *et al*. (1999) emphasized the effect of net O_3 chemical production at midlatitudes. Li *et al*. (2002) further suggested that the apparent correlation between surface observations of O_3 and ^7Be in Bermuda merely reflects the subsidence of O_3 produced in the middle troposphere, not O_3 transported from the stratosphere.

Applying global 3-D chemistry and transport models, a number of recent studies have investigated the effects of intercontinental transport on tropospheric O_3 concentrations from Asia to North America (e.g., Berntsen *et al*., 1999; Jaffe *et al*., 1999; Jacob *et al*., 1999; Bey *et al*., 2001). The effect of trans-Pacific transport is particularly large in spring

(e.g., Wang *et al*., 1998; Jacob *et al*., 1999; Mauzerall *et al*., 2000; Wild and Akimoto, 2001; Tanimoto *et al*., 2002).

We analyze here aircraft springtime observations from TOPSE (February–May 2000) (Atlas *et al*., 2003). Thirty eight science flights of the C-130 aircraft were conducted in seven deployments (1–2 weeks apart), covering a latitude range of 40° – 85 °N (Colorado to north of Thule, Greenland) and an altitude range from the surface up to 8 km. A comprehensive suite of chemical species related to the tropospheric chemistry of O_3, HOx ($OH + HO_2$), and NOx ($NO + NO_2$) were measured.

This data set is the largest available that extends over broad latitude and altitude ranges on a seasonal timescale in spring. We will attempt to address in this work the effects on tropospheric O_3 variability and trend by intercontinental transport of O_3 and its precursors from different tropospheric regions and by transport of O_3 from the stratosphere on the basis of TOPSE observations. We will apply a multivariate analysis technique in this work. As illustrated by Li *et al*. (2002) and Browell *et al*. (2003), considering multiple correlations among O_3 and other species is critical.

4.2. Methodology

In this work, we analyze the covariant factors of 14 relatively long-lived chemical tracers and potential temperature. Our hypothesis is that air masses with varying contributions to the observed O_3 variability and trend can be differentiated on the basis of the covariance of their chemical compositions and that their geographical origins can be evaluated using back trajectory calculations.

To examine the tracer profiles of various air masses, we applied two analytical methods, principal component analysis (PCA) and the positive matrix factorization (PMF) method. Covariance among the tracers is analyzed in these methods to define key components or factors that explain the variability of the data set. The PCA method (Press *et al*., 1992) with varimax rotation (Kaiser, 1958) yields comparable results as those by PMF. We find, however, that the PMF results are clearer for physical interpretations and for the separation of air masses than those by PCA because the latter lumps positively correlated tracers with negatively correlated tracers. Furthermore, the PMF results are more robust with respect to rotation. Only PMF results are discussed hereafter.

Our results cannot be directly compared to model estimated fractional contributions to tropospheric O_3 from different sources. Any correlation-based methods including the ones used here can only be used to examine the variability in the observations. However, the results from variability analysis as those to be presented here are valuable because (1) they are based on the observations and are not affected by model imperfections and (2) the resulting concentration variability is often a key parameter with regard to potential environmental impact of a given factor. In order to fully reconcile our data analysis with those from the 3-D chemistry and transport models, we will need to conduct detailed simulations with the same species and equivalent spatial and temporal resolutions as in the observations. That is beyond the scope of this work. It is also useful to note that the validity of the linear covariance assumption of PMF, PCA, and other similar analysis techniques is tested on whether the resulting factors or principal components can be physically explained. These mathematical techniques evaluate only the ‘‘likelihood’’ of a

given setup of factors or principal components; the results may be nonphysical. We choose the solution that entails the best physical explanation.

The profiles of covariant elements (factors) are identified through matrix decomposition in the PMF method (Paatero, 1997) and the detail explanation of terms and definitions are described in Chapter 1.

Among the available measurements during TOPSE, we chose 14 chemical tracers and one dynamic tracer. They are among the longer-lived tracers in the TOPSE measurements.

The tracers (other than O_3) must also have unique source or concentration distributions, so they could be used to identify the origins of air masses. The chemical tracers chosen for the analysis are O_3, total reactive nitrogen (NO_y), peroxyacetylnitrate (PAN), CO, CH_4, C_2H_2, C_3H_8, CH_3Cl, CH_3Br, C_2Cl_4, CFC-11 (CCl_3F), HCFC-141B (CH_3CCl_2F), Halon-1211 ($CBrClF_2$), and ^7Be. Total reactive nitrogen is a good tracer of air masses that intercepted large tropospheric NOx emissions or stratospheric air. PAN is a by product of NOx oxidation in the presence of >C2 hydrocarbons. It is a good tracer for tropospheric reactive nitrogen because the lifetime of >C2 hydrocarbons is too short to be transported into the stratosphere. High concentrations of PAN often indicate occurrences of tropospheric O_3 production in the past and therefore tend to correlate with high O_3 concentrations in pollution plumes. Carbon monoxide and C_2H_2 are good tracers for combustion and C_3H_8 is a good liquefied gas tracer. Methane has many sources, some of which are collocated with industrial pollution sources. Methyl chloride and CH_3Br are thought to have substantial sources from the ocean and biomass burning. Tropical terrestrial biosphere is likely a major CH_3Cl source region (Yokouchi *et al*., 2000; Lee-

Taylor *et al*., 2001). Tetrachloroethene is mainly a tracer for industrial activities. The production of CFC-11 and Halon-1211 was phased out in the developed countries at the end of 1995 and will be phased out in developing countries by 2010. HCFC-141b is a replacement for CFCs. Beryllium-7 is produced by cosmic rays and is generally used as a tracer for stratospheric air (e.g., Dibb *et al*., 2003). The other tracer we use in the analysis is potential temperature (θ), which is conserved during adiabatic processes. Potential vorticity is another useful dynamic tracer. It needs to be computed on the basis of θ and model simulated horizontal winds; the resulting temporal and spatial resolutions of this variable are much lower than the in situ measurements.

The observed values of each variable are linearly scaled to a dimensionless range of 0–1 and a small constant error estimate is assigned for equal weighting. In our analysis, no tracer should be weighted more than the other because they represent different processes. Our goal is not to define the best correlation between two specific tracers but rather to define the best tracer characteristics of air masses on the basis of the variability of all the tracers. Only data above the detection limits are used. By evaluating the error matrix E (Paatero, 1997; Lee *et al*., 1999; Paatero *et al*., 2002), we found that a range of 6–10 factors are appropriate for our data sets. We chose to use seven factors, which separate air mass tracer profiles well.

The 10 day kinematic back trajectories used in this work are calculated using the Hybrid Single-Particle Lagrangian Integrated Trajectory (HY-SPLIT) model (Draxler and Hess, 1998). The meteorological data are taken from the 'FNL' archive produced in the NOAA Air Resources Laboratory. The 6 hour archive is computed using a global data assimilation system (Kanamitsu, 1989). This type of trajectory calculation is generally

accurate for identifying large geographical features of air mass origins (Stohl, 1998). The choice of 10 days for the back trajectory calculation is somewhat arbitrary; it is long enough to separate local influence from long-range transport (Pochanart *et al*., 2001).

4.3. Results and Discussion

The photochemical environment is quite different from middle to high latitudes due in part to varying solar isolation (Wang *et al*., 2003a). Consequently, our analysis is conducted separately for these two regions. The size of our data set is limited mostly by the availability of ^{7}Be data to 137 coincident measurements of all species at midlatitudes (40° – 60°N) and 193 coincident measurements at high latitudes (60° – 85°N). Including missing ^{7}Be data by assigning large uncertainties (Polissar *et al*., 1998) led to an underweight of the ^{7}Be data, resulting in a poor O_3 -^{7}Be correlation and a large fraction of unexplained ^{7}Be variability. As demonstrated by Li *et al*. (2002) and Dibb *et al*. (2003), faithfully capturing the O_3-^{7}Be correlation is critical to resolve the contributions by various O_3 sources. Hence no missing ^{7}Be data were included. Ozone concentrations in this subset show similar probability distributions (binned by 10 ppbv) as the data set with all measured O_3. The altitude distribution is shifted to higher altitude. The fraction of data above 5 km increases from 50% of all O_3 data to 70%. The seasonal O_3 trend of the subset is in agreement with that of all O_3 measurements above 5 km. Our results are therefore biased toward the middle and upper troposphere. The analysis results are about the same when measurements below 3 km are excluded.

In order to establish the correlation between the stratospheric O_3 and ^{7}Be, we have included in our analysis data points with O_3 concentrations >100 ppbv (5% of the data

set), which are generally associated with the lower stratospheric air. To minimize the effects of these lower stratospheric data, all results shown hereafter are for data with O_3 concentrations <100 ppbv. Relative to the results with all the data included, the effects of the limited O_3 range are only apparent for O_3 and ^{7}Be and the estimated stratospheric contribution to O_3 variability is reduced. The EV percentages of O_3 in the factors change by <5%. The largest changes are seen in the explained variations of O_3 and ^{7}Be in the ^{7}Be factor, which decrease by 6 ppbv and 100–150 fCi/SCM, respectively.

4.3.1. Midlatitudes

The EV profiles of the seven factors and the total EV explained by the analysis for data with O_3 concentrations <100 ppbv are shown in Figure 4.1. The factors are named after their major components. Our analysis captures 80–95% of the observed variation of each tracer. The absolute tracer variation explained by each factor is shown in Figure 4.2. Five factors are mostly related to the sources and transport of CO, hydrocarbons, and halocarbons. The seasonal decrease of CO and hydrocarbons (e.g., Blake *et al*., 2003) is captured by the hydrocarbon factor.

Among the seven factors, only ^{7}Be and NO_y-PAN factors account for >5% of the observed variability of O_3. The tracer profile of the ^{7}Be factor is consistent with a stratospheric origin. Potential temperature and HNO_3 concentrations in the stratosphere are much higher than those in the troposphere. The latter is reflected in the signal of NO_y, which encompasses all reactive nitrogen species. The factor shows no signal of PAN, which is produced in the troposphere. On the basis of our preliminary global 3-D simulations of CH_3Cl, we believe that the signal of CH_3Cl in the ^{7}Be factor likely reflects

a general latitudinal increase of CH_3Cl toward tropics and the isentropic mixing of tropical and stratospheric air. The variability of CH_3Cl and CH_3Br explained by this factor is relatively small at 4 pptv and 0.15 pptv, respectively.

We estimate that stratospherically influenced air accounts for about 40% (14 ppbv) of the observed O_3 variation (Figures 4.1 and 4.2). An uncertainty in our analysis comes from the significant ^7Be source in the upper troposphere (e.g., Liu *et al*., 2001; Allen *et al*., 2003). However, ^7Be concentrations are much lower in the upper tropospheric than the stratosphere due to the large difference of air residence time. Allen *et al*. (2003) showed that the variability due to ^7Be produced in the stratosphere is much larger than that produced in the troposphere. It is a good tracer for the stratospheric air in the context of the observed variability during TOPSE (D. Allen, personal communication, 2002). Furthermore, an upper tropospheric ^7Be source does not necessarily imply that it will bias our results to favor the effect of stratospheric O_3. On the contrary, it could lead to an underestimate of the effect of stratospheric O_3 if the enhancement in ^7Be is not accompanied by comparable enhancement in O_3 concentrations. During TOPSE, the $\Delta O_3/\Delta^7Be$ ratio from a least squares fit for data above 6 km is about half of that for those below 6 km; the least squares fitted $\Delta O_3/\Delta^7Be$ ratio for the whole data set is about the same as that for data above 6 km.

The Bermuda O_3 -^7Be correlation analysis by Li *et al*. (2002) appears to show results that differ from our work for TOPSE. The data they used were collected near the surface during March–May 1996 in Bermuda (32°N). The season is similar but the latitude and sampling altitude are different. During TOPSE, ^7Be concentrations up to 4000 fCi/SCM were observed while the Bermuda site measurements are <250 fCi/SCM. A least squares

fit of TOPSE midlatitude observations yields a slope of 0.04 ppbv/(fCi/SCM). The correlation coefficient is 0.8, indicating that the variance of ^{7}Be explains about 65% of the O_3 variance. The variability ratio of O_3 to ^{7}Be in the ^{7}Be factor shows a similar value, 0.045 ppbv/(fCi/SCM) (Figure 4.2).

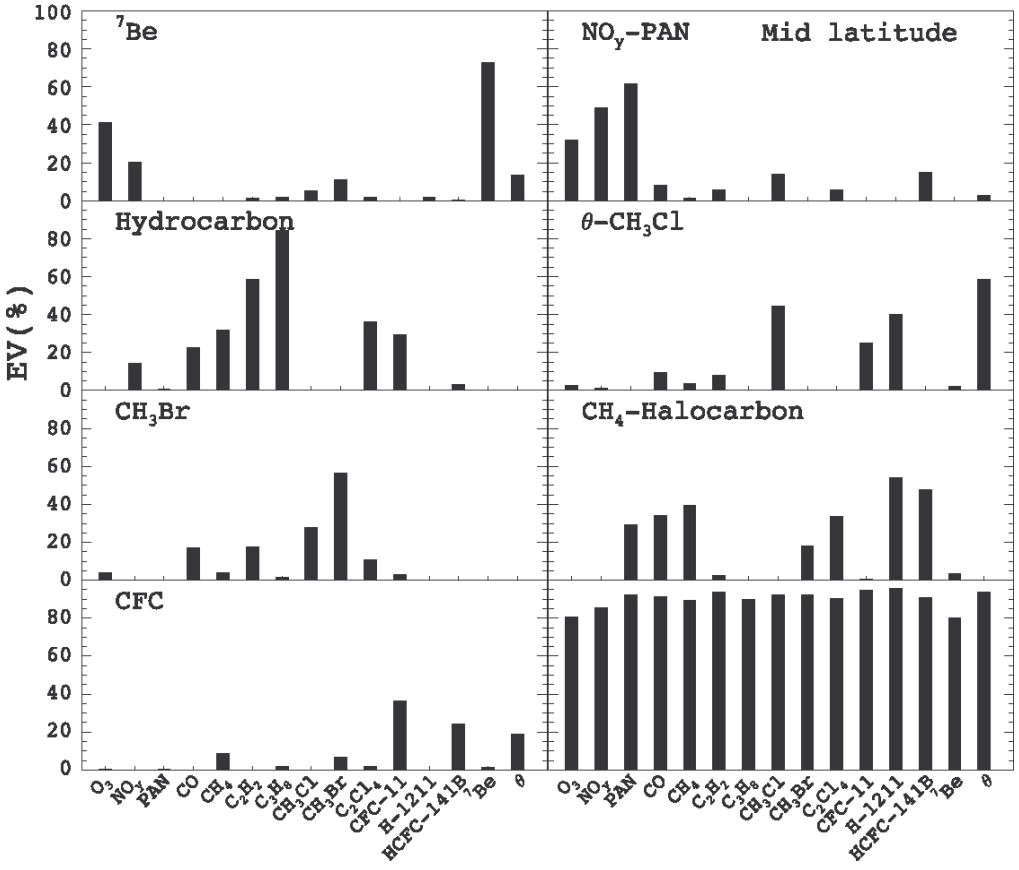


Figure 4.1. Explained variation of each tracer in the seven factors at midlatitudes (40° – 60° N). The bottom panel in the right column shows the explained variation (EV) percentages explained by all the factors. The factors are named after the tracers that have strong signals. The results are for data with O_3 concentrations < 100 ppbv.

In comparison, Li *et al*. (2002) showed a much larger slope of 0.45 ppbv/(fCi/SCM); the correlation coefficient is very low, at only 0.36. The much higher O_3 -^{7}Be slope at

Bermuda indicates that transport from the stratosphere is not the major source of O_3 variability at Bermuda. If we simply scale the $\Delta O_3/\Delta^7Be$ ratio obtained from the PMF analysis to the maximum variability of ^7Be at Bermuda (about 160 (fCi/SCM)), we would estimate a maximum stratospheric contribution of 7 ppbv to the observed O_3 variability at Bermuda during March–May 1996. Furthermore, if we assume that only 0.045 out of 0.45 ppbv/(fCi/SCM) in the observed $\Delta O_3/\Delta^7Be$ ratio can be expected from the stratospheric contribution based on our PMF analysis, the tropospheric contribution to the observed O_3 variability is about 90% at Bermuda. This result is consistent with that by Li *et al*. (2002).

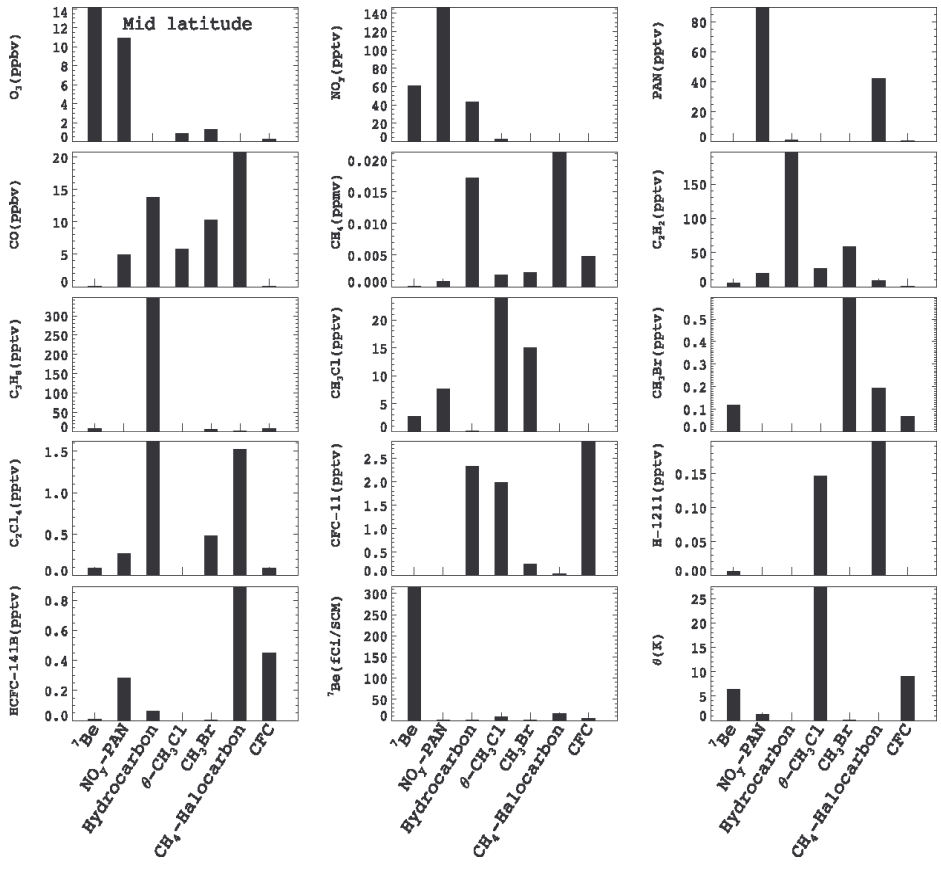


Figure 4.2. The tracer variation ([tracer] – [tracer]$_{min}$) explained by each factor. The results are for data with O_3 concentrations < 100 ppbv.

The NO_y-PAN factor reflects O_3 production within the troposphere. During the oxidation of CO and hydrocarbons, freshly emitted NO_x can be converted to HNO_3 and PAN, and O_3 is produced. The observed O_3 variation accounted for by this factor is about 30%, less than that by the ^7Be factor. There are only two main factors for PAN variability (Figure 4.2). Beside the NO_y-PAN factor, the other one is the CH_4-halocarbon factor. The latter is the most dominant factor for CO and CH_4 variability. It also contributes largely to the variability of the industrial tracer C_2Cl_4, suggesting that this factor is associated with air masses affected by industrial emissions.

There is not a significant PAN signal in the hydrocarbon factor, which has large CO, CH_4, and C_2Cl_4 signals because the seasonal trend of this factor is opposite to that of PAN. The lack of NO_y signal in the CH_4-halocarbon factor, despite its large PAN signal, may reflect the different seasonal trends of NO_y and PAN during TOPSE (Wang *et al.*, 2003a). The sum of NO_y components including PAN and HNO_3 is generally higher than the observed NO_y and the former has a stronger seasonal trend; the reason for this discrepancy is unclear (Wang *et al.*, 2003a). A least squares fit of PAN and NO_y shows that the variance of PAN explains about 30% of that of NO_y.

Figure 4.3 shows the 10 day back trajectories for data points with high factor scores for the ^7Be and NO_y-PAN factors. Both are associated with the westerlies at northern midlatitudes in spring. The trajectories related to the ^7Be factor tend to extend farther west into Europe whereas those related to the NO_y-PAN factor tend to cluster over the western Pacific. The back trajectory altitude of the ^7Be factor resides mostly above 6 km, while those related to NO_y-PAN factor reside more likely at 2–6 km. Very few back trajectories initiated at 0–2 km or south of 30°N.

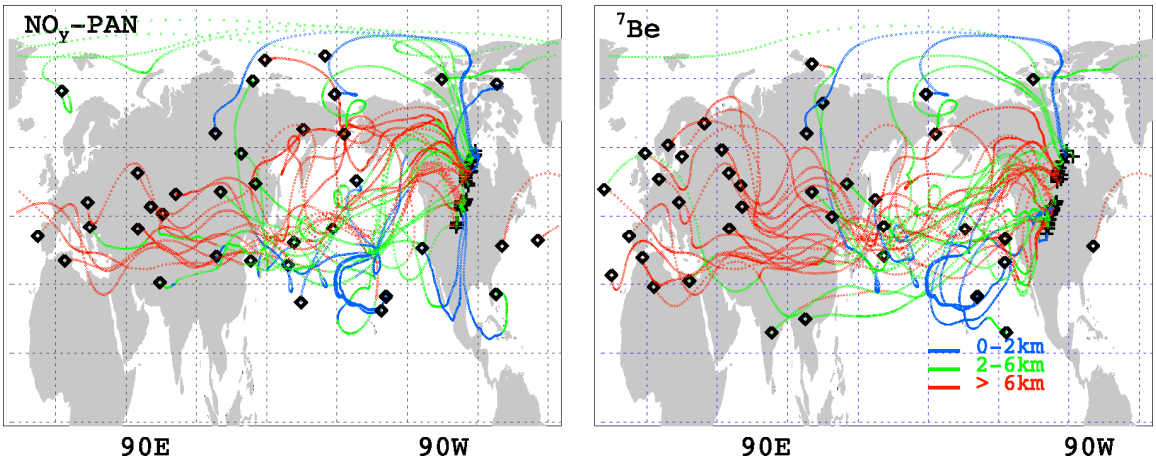


Figure 4.3. Ten day back trajectories of data points that rank in the top 30% by factor score for the NO_y – PAN and ^7Be factors at midlatitudes. The crosses show the locations of measurements. The back trajectory location every hour is represented by a dot, and the end points is marked by a diamond. The results are for data with O_3 concentrations < 100 ppbv.

Midlatitude emissions of O_3 precursors from East Asia, after being lifted out of the boundary layer, could have contributed significantly to the O_3 enhancement in the NO_y-PAN factor. Pollutions from the Middle East and southern Europe, lifted by convection to high altitudes (>6 km), could also contribute to the enhancement.

The observed tropospheric O_3 increase at midlatitudes can now be attributed to different factors. We conduct the least squares fits for the variability of observed O_3 and that explained by each factor (Figure 4.4). We analyze the linear seasonal trend because it reasonably represents the increase of monthly median O_3 variation and facilitates the evaluation of the contribution from different factors. The observed seasonal O_3 increase is 5.4 ppbv/month while that from the contribution of all factors is 6.3 ppbv/month.

The seasonal increase contributed by the NO_y-PAN factor is largest at 3.5 ppbv/month and that from the ^7Be factor is at 2.5 ppbv/month. Contributions from the other factors

are negligible. Photochemistry within the troposphere therefore appears to be more important to the seasonal O_3 increase than transport from the stratosphere.

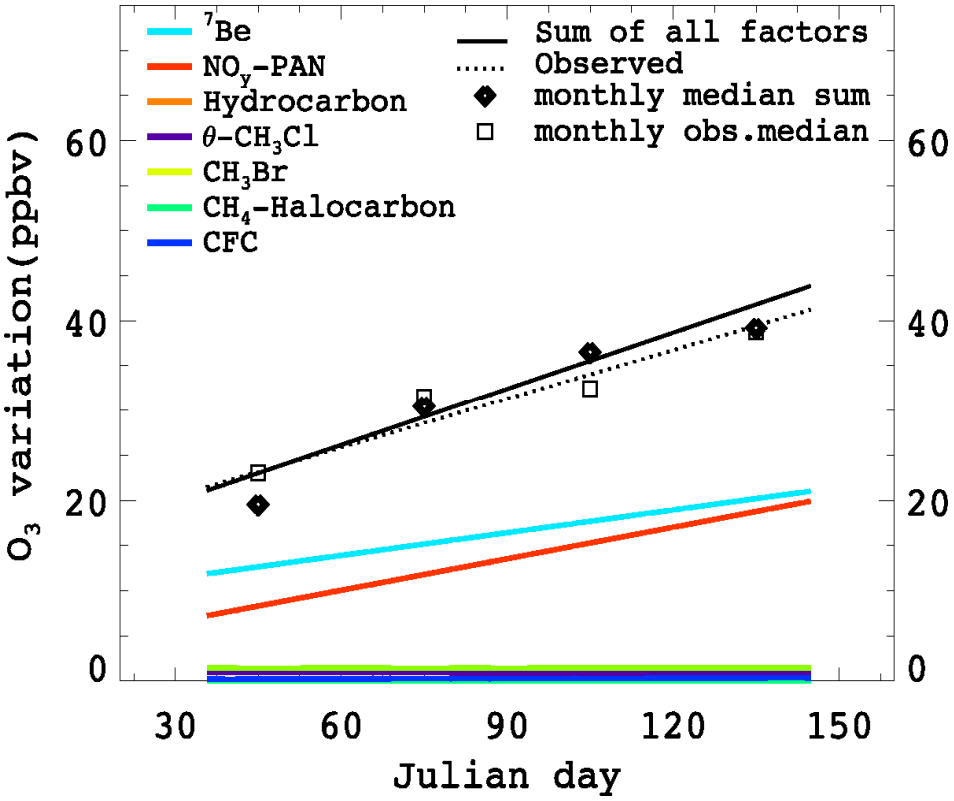


Figure 4.4. Least squares fits of O_3 variation ($[O_3] - [O_3]$min) as a function of Julian day for the observed and positive matrix factorization-fitted O_3 and that explained by each factor at midlatitudes. The monthly medians for the O_3 variations from the observations and the sum of all factors are also shown. The minimum O_3 concentration in the data set is 27 ppbv. The results are for data with O_3 concentrations < 100 ppbv.

4.3.2. High Latitudes

At high latitudes, the ^{7}Be factor is still the largest contributor to the observed O_3 variability at about 35% (Figure 4.5.). The $\Delta O_3/\Delta^7$Be ratio in this factor is 0.03

fCi/SCM/ppbv (Figure 4.6), lower than at midlatitudes. The lower ratio is entirely due to higher variability in ^{7}Be at high latitudes, which is consistent with higher ^{7}Be sources near the poles.

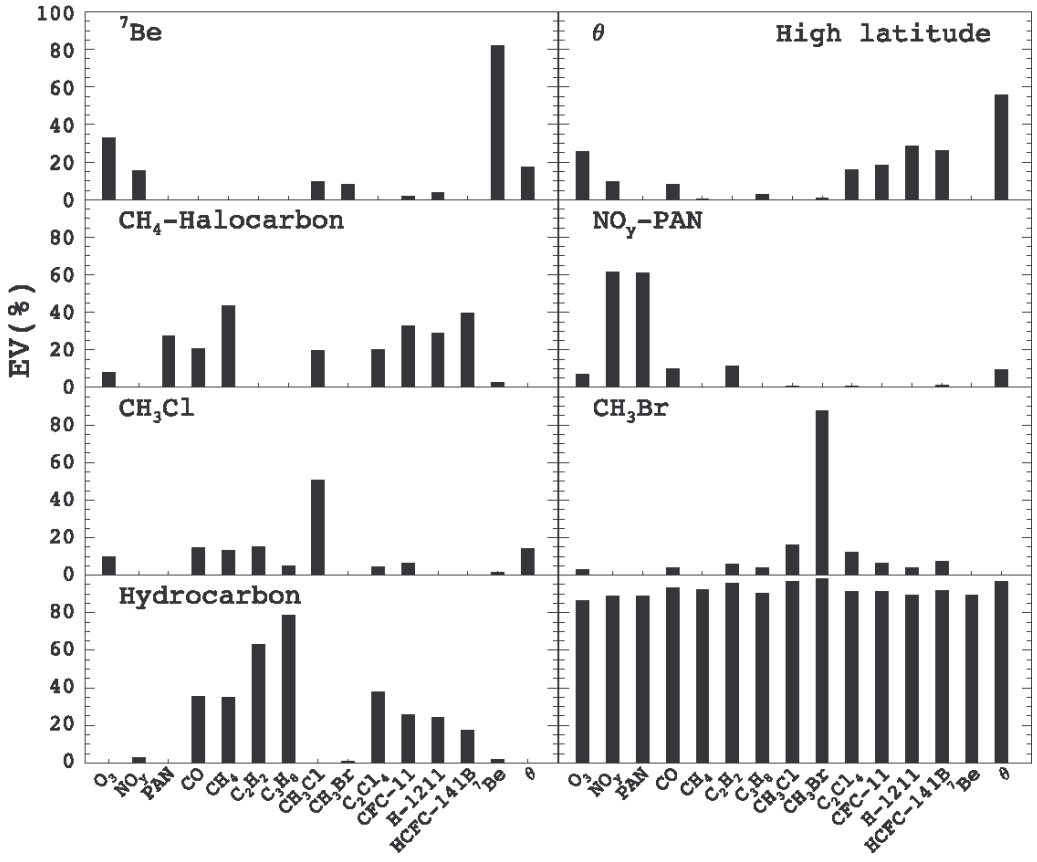


Figure 4.5. Same as Figure 4.1 but for high latitudes (60° – 85°N).

The θ factor is the other one that has a large contribution (25%) of the observed O_3 variability. It also explains 60% and 20 – 30% of the variability in θ and many halocarbons, respectively. The NO_y-PAN factor, in contrast, accounts for only about 5% of the observed O_3 variability.

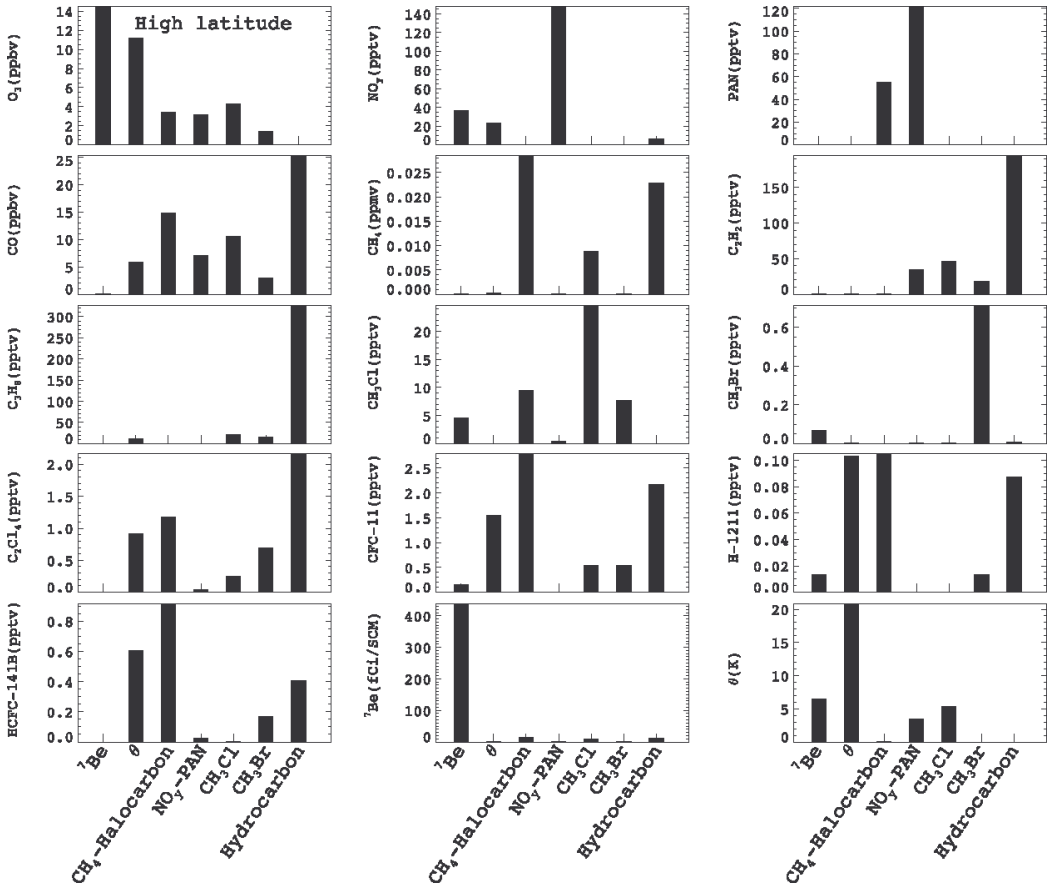


Figure 4.6. Same as Figure 4.2 but for high latitudes.

The 10 day back trajectories for data points with high factor scores in the ^{7}Be and θ factors at high latitudes are quite different (Figure 4.7). A large fraction of data contributing to the ^{7}Be factor circulated within the polar region north of 60°N. These back trajectories reside generally above 6 km. In contrast, most of the data for the θ factor came from regions south of 60°N carried by the strong westerlies and were transported northward in the western United States. The related back trajectories more likely reside below 6 km. Higher O_3 concentrations appear to be transported in the middle troposphere from the lower latitudes into high latitudes. Despite low in situ chemical

production of O_3 (Wang *et al.*, 2003a), photochemical production of O_3 within the troposphere, through long range transport from lower latitudes where photochemistry is active, still contributes significantly to the observed O_3 variability at high latitudes.

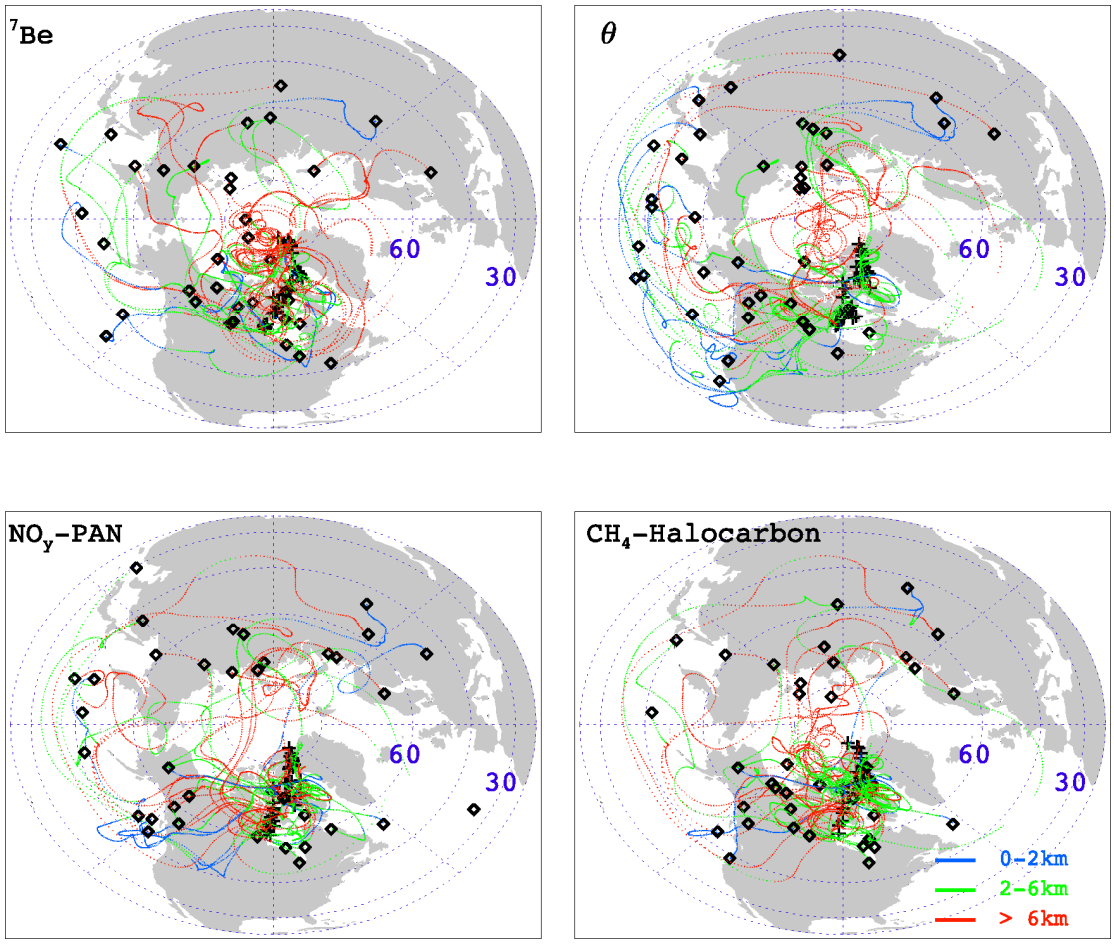


Figure 4.7. Same as Figure 4.3 but for the ^{7}Be, θ, NO$_y$-PAN, and CH$_4$-halocarbon factors at high latitudes.

These two factors also account for much of the explained variability in θ. The gradient of θ is much stronger from the stratosphere into the troposphere as compared to that from middle to high latitudes in the troposphere (e.g., Holton *et al.*, 1995); the former is reflected in the vertical stability in the stratosphere and the latter is limited by

tropospheric baroclinic adjustment processes. Similarly we find that the θ -CH_3Cl factor that explains most of the variability of θ at midlatitudes (Figure 4.1) is related to transport from the subtropics into midlatitudes (not shown) and that it is associated with transport of CH_3Cl into the region. The larger contribution of latitudinal transport to the observed variability of θ across a much smaller gradient compared to transport from the stratosphere reflects much stronger latitudinal mixing in the northern hemisphere.

Although the ^7Be and θ factors account for about 60% of the observed O_3 variability at high latitudes, their contribution to the seasonal trend of tropospheric O_3 is much smaller (Figure 4.8). The observed seasonal trend for tropospheric O_3 at high latitudes is 5.5 ppbv/month; the trend explained by all seven factors is 4.6 ppbv/month. The ^7Be and θ factors account for 0.8 and 0.6 ppbv/month, respectively.

The bulk of the seasonal O_3 increase is attributed to the NO_y-PAN and CH_4-halocarbon factors, accounting for 1.3 ppbv/month and 1.7 ppbv/month, respectively. These two factors also explain about 90% of the variability in observed PAN concentrations (Figures 4.5 and 4.6). PAN is of tropospheric origin. These two factors therefore signify the effects of tropospheric production on the seasonal increase of tropospheric O_3.

The 10 day back trajectories for data points with high factor scores in the NO_y-PAN factor (Figure 4.7) show two primary pathways of transport via either midlatitude westerlies over the Pacific or across the polar region. East Asia and northern Europe are the likely source regions for the transported reactive nitrogen. Comparing to the transport of reactive nitrogen at midlatitudes (Figure 4.3), the transport into the high latitudes tends to occur more below 6 km and is more likely to be affected by surface emissions.

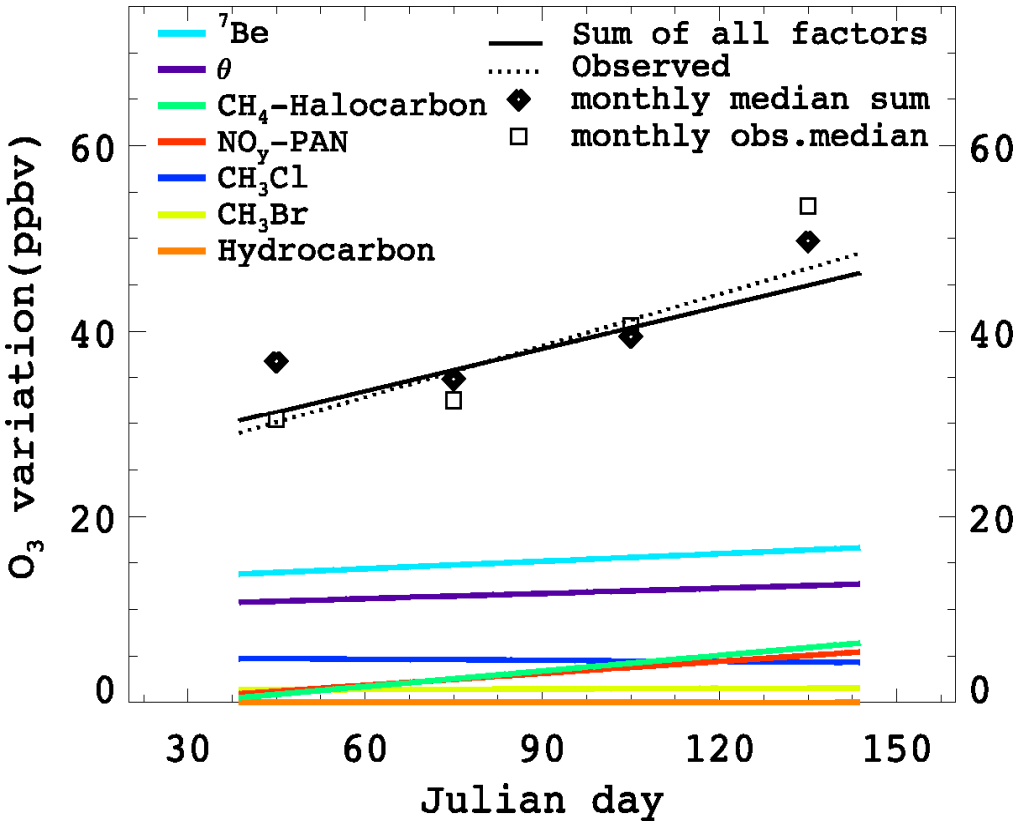


Figure 4.8. Same as Figure 4.4 but for high latitudes. The minimum O_3 concentration in the data set is 26 ppbv.

Air masses contributing to the CH_4-halocarbon factor tend to be more stagnant than those to the NO_y-PAN factor with a large fraction of air circulating over Canada in the last 10 days. The tracer profile of the CH_4-halocarbon factor also reflects the origins of air masses. Among the seven factors, two account for about 80% of the observed CH_4 variability. One is the CH_4-halocarbon factor. The other factor, which accounts for about 35% of the observed CH_4 variability, is associated with the seasonal decrease of hydrocarbons as defined largely by C_2H_2 and C_3H_8 (Figure 4.5). The strong CH_4 signal in

the CH_4-halocarbon factor likely reflects natural gas leaks from Russia, which has a large reservoir and active production of natural gas (van Meurs, 1997).

The CH_4-halocarbon and θ factors have large halocarbon signatures (Figures 4.5 and 3.6). These two factors and the hydrocarbon factor explain most of the observed variability in these species. Production of CFC-11 and Halon-1211 has been banned in developed countries. The halocarbon signals in the cross-Pacific transport as illustrated by the θ factor likely reflect the release of these halocarbons from China. The halocarbon signals in the CH_4-halocarbon factor appear to suggest continuing release of halocarbons (including CFC-11 and Halon-1211) from Russia.

Our results on the stratospheric contribution to the O_3 variability and seasonal trend are generally consistent with those derived from the global 3-D simulation by Allen *et al*. (2003). The two results cannot be compared directly because we can only examine the variability and trend using our methodology while Allen *et al*. (2003)could estimate the stratospheric contribution to O_3 concentrations. Dibb *et al*. (2003), by analyzing the O_3-^{7}Be correlation, found that highlatitude O_3 from the stratosphere does not contribute to the seasonal O_3 increase; we draw a similar conclusion in our analysis. They estimated that >85% of high-latitude O_3 is related to the stratospheric source by further interpreting the ‘‘residual O_3’’ from the intercept values of O_3 -^{7}Be least squares fitting. This estimate is larger than our estimates, although we have attempted to resolve only the variability. Browell *et al*. (2003) suggests that >80% of the O_3 trend is due to tropospheric photochemistry on the basis of O_3 correlations with observed aerosol concentrations and model computed potential vorticity. Our result at high latitudes agrees well with their estimate. At midlatitudes, we estimate a much higher stratospheric contribution (about

40%) to the observed O_3 trend. It is unclear if their estimate would be closer to ours if only observations at midlatitudes (rather than the whole data set) are considered in their analysis.

4.4. Conclusions

PMF and PCA analysis techniques are applied to analyze the TOPSE observations to investigate the various factors contributing to the observed variability of tropospheric O_3 and its seasonal trend observed during spring 2000 at middle and high latitudes. A suite of 14 chemical tracers and one dynamic tracer were used to define the chemical characteristics of the air masses. The coincident data of 15 tracers we used in the analysis are a small subset of all TOPSE measurements due largely to the availability of ^7Be measurements. As a result, our results are biased toward the middle and upper troposphere. We find no bias in the probability distribution of O_3 concentrations (binned by 10 ppbv) in the subset, but a larger fraction of measurements, 70% vs. 50% for the whole data set, was taken at 5– 8 km. The seasonal O_3 trend in the subset is in agreement with all O_3 data above 5 km.

We find that the PMF results are clearer for physical interpretation than those of PCA. Computed 10 day back trajectories contributing to each factor are examined to define the geographical origins of the air masses. The chemical characteristics from factor analysis and the associated back trajectories provide complimentary information that is in general consistent with our current understanding of the source distributions of chemical and dynamical tracers. While it cannot offer the depth of information that can be extracted from 3-D model simulations, the analysis provides important observation-

based constraints on the controlling factors for springtime O_3 variability and its seasonal trend at northern middle and high latitudes.

The intercontinental nature of pollution transport in spring is clearly evident in our results. At midlatitudes, the timescale of transport of pollutants from Eurasia to North America is only 1–2 weeks by midlatitude westerlies. Coupled with a favorable photochemical environment, O_3 production by reactive nitrogen emitted from Eurasia (with possible contributions from North America) is a major factor leading to the observed O_3 variability and seasonal trend during TOPSE. We estimate that this source, related to the PAN-NOy factor, explains an O_3 enhancement of 11 ppbv and a seasonal trend of 3.5 ppbv/month.

At high latitudes, the effect of intercontinental transport becomes even more prominent in part because springtime photochemistry in the region is relatively slow. It is manifested in the transport of O_3 and its precursors. The dominant tropospheric factor affecting the observed O_3 variability is related to the transport of potential temperature (θ factor) from lower latitudes, where the high availability of photons and anthropogenic emissions result in a general ‘‘background’’ O_3 enhancement compared to high latitudes. We estimate that it contributes to an O_3 enhancement of about 11 ppbv but only 0.6 ppbv/month in the seasonal trend. In contrast, the PAN-NOy and CH_4-halocarbon factors, which explain all the variability of PAN (120 and 60 pptv, respectively), contribute respectively 1.3 and 1.7 ppbv/month to the seasonal trend even though their contribution to the O_3 variability is about 4 ppbv each. Therefore transport of O_3 from lower latitudes, which is poor in reactive nitrogen, tends to dominate O_3 variability while transport of reactive nitrogen-rich air masses tends to dominate the seasonal trend. Cross-Pacific

transport is important in all cases; cross-Arctic transport is also important for reactive nitrogen-rich air masses. In addition, the halocarbon signals in the CH_4-halocarbon factor appear to indicate continuing release of halocarbons (including CFC-11 and Halon-1211) from Russia.

Transport from the stratosphere, diagnosed through the ^7Be factor, is also important. It is manifested mostly above 6 km while transport of tropospheric O_3 and its precursors is more likely to be important at 2 – 6 km, indicating that the stratospheric influence decreases with altitude. We estimate that it contributes to an O_3 enhancement of about 14 ppbv (35 – 40%) for data with O_3 concentrations <100 ppbv at middle and high latitudes. There is a large latitudinal difference in its contribution to the seasonal trends. It is 2.5 ppbv/month (40%) at midlatitudes but only 0.8 ppbv/month (<20%) at high latitudes. The difference probably reflects increasing activity of the more vigous dynamic systems at midlatitudes in spring.

CHAPTER V

EVALUATION OF MODEL-SIMULATED SOURCE CONTRIBUTIONS TO TROPOSPHERIC OZONE WITH AIRCRAFT OBSERVATIONS IN THE FACTOR-PROJECTED SPACE[1]

5.1 Introduction

As explained in previous chapters, troposphere O_3 has important environmental consequences. The sources of tropospheric O_3 consist of either photochemical production within the troposphere or transport from the stratosphere. Many studies have tried to investigate tropospheric O_3. O_3 concentration at northern high and mid latitudes tends to peak in late spring (April and May) in the lower troposphere and in early/middle summer (May – July) in the middle troposphere (e.g., Logan, 1985). Penkett and Brice (1986) and Liu *et al.* (1987) suggested that the increase is driven by photochemical production in the troposphere.

A number of studies using 3-D chemical transport models (e.g., Berntsen *et al*., 1999; Jaffe *et al.*, 1999; Jacob *et al.*, 1999; Bey *et al.,* 2001) have investigated the effect of intercontinental transport on tropospheric O_3 concentrations from Asia to North America. The effect of trans-Pacific transport is particularly noticeable in the spring (e.g., Wang *et al.*, 1998, 2006; Jacob *et al.*, 1999; Mauzerall *et al.*, 2000; Wild and Akimoto, 2001; Tanimoto *et al*., 2002). However, other studies based on the observed correlations between O_3 and ^7Be attributed this trend to the effect of the stratospheric input (e.g., Oltmans and Levy, 1992; Dibb *et al.*, 1994).

[1] This chapter is for "Evaluation of model-simulated source contributions to tropospheric ozone with aircraft observations in the factorp-projected space", prepared for the submission to *Journal of Geophysical Research* in 2006. Authors are **Changsub Shim**, Yuhang Wang, and Yasuko Yoshida.

Tropospheric Ozone Production about the Spring Equinox (TOPSE) aircraft observations were conducted (February – May 2000) (Atlas *et al.,* 2003). Previously, an advanced factor analysis using positive matrix factorization (PMF) with TOPSE measurements found that an increasing seasonal trend of springtime O_3 at northern mid and high latitudes is most responsible for the intercontinental transport of polluted air masses, even though stratospheric influence is the largest contributor to O_3 variability (Wang *et al.*, 2003b). The PMF analysis with measurements is applied because it is often advantageous over the study of the 3-D chemical transport model when the source types, strength, distributions and transport are not clearly understood.

Here, we apply the factor analysis to a suit of simulated data from a 3-D chemical transport model as well as aircraft measurements in order to evaluate the simulated contributions to tropospheric O_3 in the projected factor space with the observations. The simultaneous comparisons between observed and simulated chemical tracers in the factor-projected space can capture the possible model problems, which is useful for the correction of the model parameterizations (Liu *et al*., 2006).

Taking advantage of the analysis, we compared the 3-D global chemical transport model (GEOS-Chem) simulated tracers with observed tracers from two aircraft missions, TOPSE and TRAnsport of Chemical Evolution over the Pacific (TRACE-P, March – April 2001) in the projected factor space. The primary object of this study is to evaluate model performance for tropospheric O_3 in the Northern Hemisphere, particularly focusing on the contributions by photochemical productions followed by surface emissions and transport within the troposphere and from the stratosphere.

We describe the data selections from TOPSE and TRACE-P and corresponding GEOS-Chem simulations in section 5.2.1. The PMF method for this analysis is explained in section 5.2.2. The results for model evaluation for TOPSE and TRACE-P are presented in section 5.3, followed by conclusions in section 5.4.

5.2 Methodology

5.2.1 Measurements and GEOS-Chem simulations

In this study, we analyze relatively long-lived chemical tracers including O_3, total reactive nitrogen (NO_y), peroxyacetylnitrate (PAN), CO, C_3H_8, CH_3Cl, and Beryllium-7 (^7Be) and one dynamic tracer (potential temperature). Those tracers other than O_3 generally have specific source characteristics. NOy is a good tracer for large tropospheric NOx emissions or stratospheric air masses. PAN can be a proxy of aged tropospheric O_3 productions and can be associated with polluted air masses because its production occurs with $>C_2$ hydrocarbon oxidation and its lifetime varies significantly with temperature (altitude). CO usually indicates combustions and C_3H_8 is a good liquefied gas tracer. CH_3Cl has its possible sources from terrestrial biosphere and biomass burning (Yoshida *et al.,* 2004). ^7Be is produced mainly by cosmic rays in the stratosphere and upper troposphere and is generally used as a tracer for stratospheric air mass (Dibb *et al.*, 2003). Potential temperature is a useful dynamic tracer due to its conservation property with adiabatic processes. The analytical approach for the observed species is similar to the work by Wang *et al*.(2003b), but the selected set of chemicals has a smaller number of tracers due to the limited availability of GEOS-Chem tracers; the resulting discrepancies with the previous work by Wang *et al* (2003b) will be discussed in section 5.3.

The photochemical and dynamic environments vary dramatically with latitudes. We separate the analysis regions to low, mid, and high latitudes. The TOPSE measurement data set are over mid (40 – 60°N, 87 – 104°W) and high latitudes (60 – 85°N, 61 – 94°W). We consider only coincident measurements, which are mostly limited by availability of ^{7}Be measurements (144 for mid latitudes and 200 for high latitudes). We exclude missing data. Including large amounts of missing data through assigning a large uncertainty to these data leads to the failure of explaining the relationship between ^{7}Be and O_3 (Wang *et al*., 2003b). The ^{7}Be and O_3 correlation is critical for analyzing the effect of stratospheric transport. The selected data has a bias towards high altitudes of 5 – 8 km (~70 % of the data), which led our analysis region toward the middle and upper troposphere.

The TRACE-P measurements data set has mid latitudes (30 – 45° N, 125 – 240 °E, 65 data) and low latitudes (15 – 30°N, 120 – 205°E, 78 data). With the same data selection criteria, the selected data also have a bias towards 7 – 12 km (40 – 50 % of the data) due to the availability of ^{7}Be measurements. Figure 5.1 shows the measurement regions during TOPSE and TRACE-P.

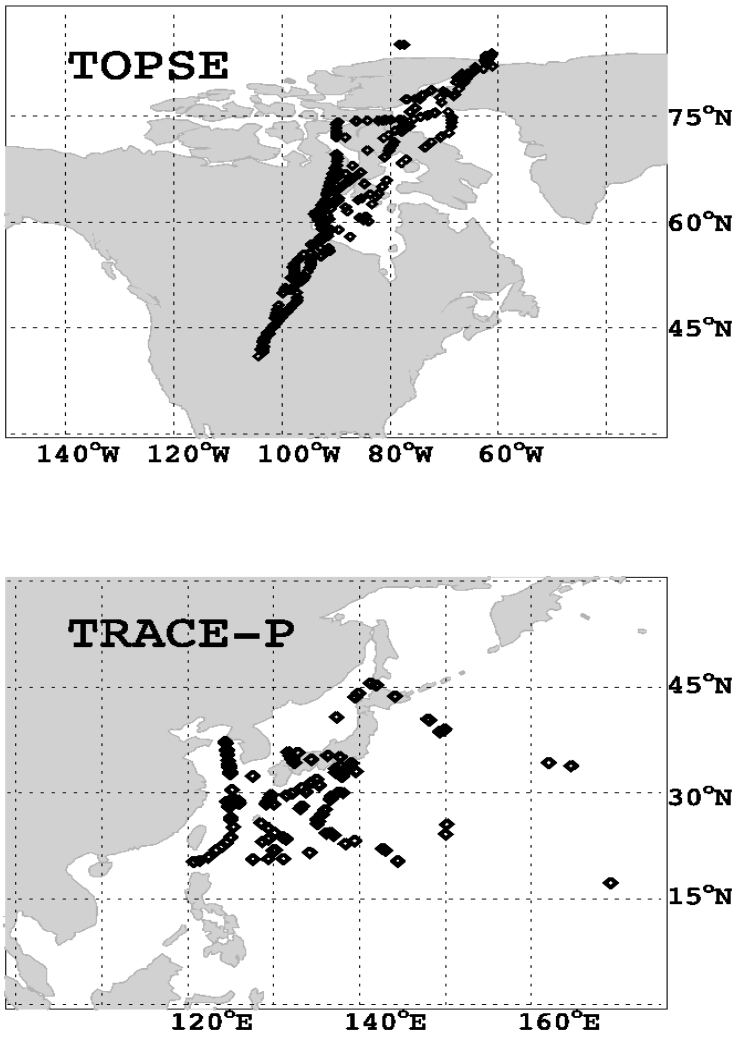


Figure 5.1. Locations of aircraft measurements during TOPSE (top) and TRACE-P (bottom).

GEOS-Chem is a global 3-D chemical transport model driven by assimilated meteorological data from the Global Modeling Assimilation Office (GMAO) (Schubert *et al.*, 1993). The 3-D meteorological fields are updated every six hours, and the surface fields and mixing depths are updated every three hours. We use version 7.24 here with a horizontal resolution of 2°×2.5° and 30 vertical layers (GEOS-3). GEOS-Chem includes

a comprehensive tropospheric O_3-NO_x-VOC chemical mechanism (Bey *et al*., 2001), which includes the oxidation mechanisms of 6 VOCs (ethane, propane, lumped $>C_3$ alkanes, lumped $>C_2$ alkenes, isoprene, and terpenes). For standard simulations, the model was spun up for one year. The GEOS-Chem simulations for the selected five tracers and one dynamic tracer (O_3, NO_y, PAN, CO, C_3H_8, potential temperature) are sampled at the same times and locations as the aircraft measurements. The simulated total reactive nitrogen (NO_y) is estimated by sum of simulated NO_x, HNO_3 (nitric acid), HNO_4 (pernitric acid), PAN, and N_2O_5 (dinitrogen pentoxide).

We follow Liu *et al*. (2001, 2004) in ^{7}Be simulations. The ^{7}Be source in GEOS-Chem is taken from Lal and Peters (1967) as a function of altitude and latitude and ~ 70% of ^{7}Be is emitted in the stratosphere. The seasonal and longitudinal dependence of ^{7}Be productions is very small and not considered. The major sink of atmospheric ^{7}Be is by wet deposition; the model considers scavenging in convective updrafts as well as first-order rainout and washout from both convective and large-scale precipitation (Liu *et al.,* 2001). Liu *et al*. (2004) reduced the ^{7}Be flux from the stratosphere by uniformly scaling down only stratospheric ^{7}Be source by a factor of 3. This simulation was first spun up for one year as well.

For model simulated CH_3Cl, we used the GEOS-Chem results conducted by Yoshida *et al*. (2004). Contributions from the six sources (pseudo-biogenic, oceanic, biomass burning, incineration/industrial, salt mash and wet land) are considered. The model results are evaluated extensively with surface and aircraft measurements; the model simulations are usually in good agreement with measurements in the northern

hemisphere. The results of the independent simulations for ^{7}Be and CH_3Cl are also sampled along aircraft flight tracks..

In order to investigate the stratospheric O_3 contributions in the model, we conducted tagged O_3 simulations to track the fraction of O_3 transported from the stratospehre (Liu *et al.,* 2002). Photochemistry is considered in the simulations by taking archived O_3 production and loss rates from the GEOS-Chem standard simulations on a daily basis. In this manner, when project simulated O_3 variability in the factor space, we can examine the fractional contribution from the stratosphere as compared to tropospheric production (section 5.3) in each factor.

5.2.2 PMF applications

The PMF method (Paatero and Tapper, 1994) explores factor categorization through the covariant structures of observed or simulated chemical and dynamical parameters (Paatero, 1997; Wang *et al*., 2003b; Liu *et al*., 2005). The PMF method basically uses the principal component analysis (PCA) scheme, but it estimates the error in the data to provide optimum data point scaling and permits a better treatment of missing and below detection limit values.

The data matrix *X* of *m* measurements by *n* tracers are decomposed in PMF analysis for *p* factors as

$$X = GF + E \tag{5.1}$$

Or

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij} \tag{5.2}$$

$$i = 1, \ldots, m; j = 1, \ldots, n; k = 1, \ldots, p.$$

where the *m* by *p* matrix G is the mass contributions of k^{th} factor to i^{th} sample (factor score), the *p* by *n* matrix F is the gravimetric average contributions of k^{th} factor to j^{th} chemical species (factor loadings), and the *m* by *n* matrix E is the error. We also use the explained variation (EV)

$$EV_{kj} = \sum_{i=1}^{m} | g_{ik} f_{kj} | / \left[\sum_{i=1}^{m} \left(\sum_{k=1}^{p} | g_{ik} f_{kj} | + | e_{ij} | \right) \right] \qquad (5.3)$$

to define the relative contributions of each factor to chemical species since the mixing ratios of different compounds are directly comparable.

As in the work by Wang *et al*. (2003a), the values of tracers are linearly scaled to a range of 0 – 1 and assigned uniformly small uncertainty for the dataset because we try to investigate the variatilities of all the tracers simultaneously and characterize the air masses. The number of factors is determined by evaluating the error matrix E (Paatero, 1997; Paatero *et al.*, 2002), and the PMF results are inspected if the characteristics of factors are clearly resolved. Based on these criteria, PMF resolved five factors for TOPSE and four factors for TRACE-P in both observed and simulated datasets.

Matrix rotation often is necessary for factor separation (Paatero *et al.*, 2002), but in general, the PMF results presented here are insensitive to rotation. Compared to direct evaluations of model results, the coherent tracer profiles of factors provide additional information to evaluate the model (e.g., Liu *et al.,* 2005). In the context of this work, we evaluate how the simulated contributions to tropospheric ozone from different processes compare to the contributions derived from observations. For this purpose, the positive

factor contributions and profiles by PMF analysis (Paatero, 1997) are important for the interpretation of the results.

5.3 Results and Discussion

5.3.1 TOPSE

As mentioned in section 5.2.1, the TOPSE results are biased toward the middle and upper troposphere. In order to capture the correlation between the stratospheric O_3 and ^7Be by PMF, we have included the data points that have O_3 concentration >100 ppbv (5% of the data set), which are generally associated with the lower stratospheric air. When analyzing the results, however, we only data points with O_3 < 100 ppbv to minimize the effect of these lower stratospheric data (Wang *et al.,* 2003b). The simulated O_3 do not exceed 100 ppbv in the selected data sets. The comparison of tracers' mean concentrations between the observation and simulation are shown in Table 5.1.

Table 5.1. The comparison of tracers' mean concentrations between observation and model.[1]

Tracers	O_3		NO_y		PAN		CO		C_3H_8		CH_3Cl		^7Be		θ	
	O[2]	M[2]	O	M	O	M	O	M	O	M	O	M	O	M	O	M
TOPSE(M)[3]	56	54.3	291	377	159	183	135	135	419	349	572	585	435	234	297	305
TOPSE(H)	65.2	54	385	396	249	220	145	144	472	418	565	575	528	233	286	293
TRACE-P(M)	64.7	61.8	1035	1002	507	379	177	186	465	626	570	626	721	252	304	309
TRACE-P(L)	57	48.7	363	668	194	171	140	160	208	338	574	631	498	178	321	323

[1]The values here are estimated with the measurement only O_3 < 100 ppbv.
The units are pptv except for O_3 (ppbv), CO (ppbv), ^7Be (fCi/SCM), and θ (K).
[2]O and M denote "observations" and "model" respectively.
[3]M, H,and L denote "midlatitudes", "high latitudes", and "low latitudes" respectively.

5.2.1.1. TOPSE at mid latitudes

PMF derived EV profiles from the observed and simulated datasets for TOPSE mid latitudes are shown in Figure 5.2. In order to quantitatively compare EV profiles between observations and model results while taking into account that the simulated variability can be quite different from the observations, we also show scaled EV profiles, which are scaled by the ratio of observed to simulated tracer variability. Even though the resulting factor profiles are not exactly same between the observations and simulations, due to the different structure of covariance between tracers, five comparable factors are resolved, reflecting fairly good simulations by GEOS-Chem (Figure 5.2). Each factor is named after the tracers that show the largest variability (^{7}Be, θ, CH_3Cl, NO_y/PAN, and Hydrocarbons).

The ^{7}Be factor in Figure 5.2 shows the largest variability of ^{7}Be for both observations and model (60% and 237 fCi/SCM for the observations, 85% and 188 fCi/SCM for the model), indicating the air masses from a stratospheric origin. The stratospheric O_3 fraction from tagged O_3 simulation in this factor show ~75% of the stratospheric origin, implying that a small fraction of the tropospheric ^{7}Be origin (~25%) is contained in this factor, which also explains the small contaminations by PAN. The ^{7}Be factor is associated with the largest O_3 variability at mid latitudes (40% and 20 ppbv for the observations, 58% and 14 ppbv for the model). A notable underestimation in the simulated ^{7}Be mean concentration is found in Table 5.1 (435 and 234 fCi/SCM, respectively).

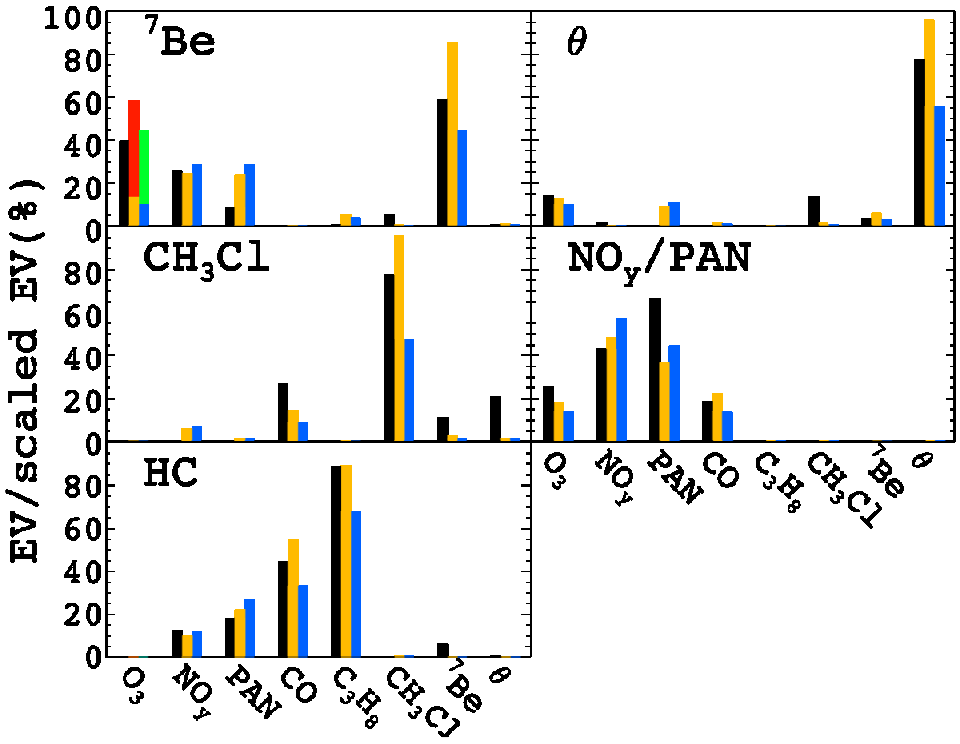


Figure 5.2. Explained variation (%, defined in section 5.2.2) of each observed tracer in the five factors at mid latitudes (40 – 60°N, left black bars). The factors are named after the tracers that have strong signals. Explained variation of each simulated tracer is compared (central orange bars). Simulated EVs are scaled by corresponding variability of observed tracer ([tracer] – [tracer]$_{min}$, right blue bars). The red and green fractions on simulated O_3 denote the stratospheric contribution from tagged O_3 simulation. The results are for data with O_3 concentrations < 100 ppbv.

As described in section 5.2, the EVs in Figure 5.2 are for data with O_3 concentration less than 100 ppbv. The ^{7}Be factor is sensitive to the O_3 – ^{7}Be correlation, and the scaled EVs are significantly changed if the lower stratospheric data (O_3 > 100 ppbv) are included (Figure 5.3). The higher O_3 variability is shown in Figure 5.3, and the scaled EV of simulated O_3 is lower due to larger observed O_3 variability (maximum observed O_3 of 217 ppbv). The lower variability of the simulated ^{7}Be is shown as well with the same reason (Figure 5.3).

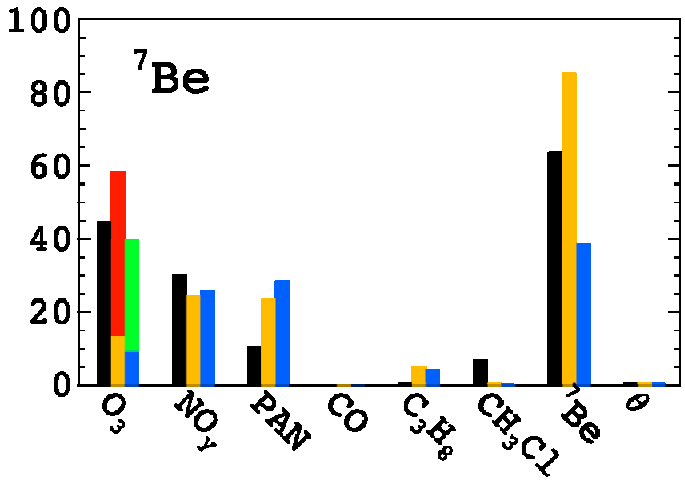


Figure 5.3. Same as Figure 5.2, but the results are for all the data including those with O_3 > 100 ppbv. Only ^{7}Be factor is shown since other factors are same.

We examine the factor correlations with latitudes, altitude, C_2H_6/C_3H_8 ratio, and CO in order to further investigate the factor characteristics (Table 5.2). The higher C_2H_6/C_3H_8 ratio reflects photochemically aged air masses (Wang and Zeng, 2004). The positive correlations of the ^{7}Be factor with altitude (r = 0.42 and 0.63 for the observed and simulated datasets, respectively) and the negative correlations with latitude (r = -0.26 and -0.37) indicate that stratosphere-troposphere exchange is more active at lower latitudes in 40-60° N region.

The potential temperature (θ) factor has large variability of θ (77% and 36.4K for the observations, 95% and 26.8 K for the model). It explains 14% of observed O_3 variability (3.6 ppbv) and 12.8% of simulated O_3 variability (3.1 ppbv). The significantly negative factor correlations with latitude (r = -0.59 and -0.51, respectively) and CO (r = -0.64 and -0.52, respectively), and positive correlation with altitudes and C_2H_6/C_3H_8 ratio (r = 0.61

and 0.43 for the observed and simulated datasets, respectively) imply that this factor is likely associated with intercontinental long-range transport of O_3 from lower latitudes, which is consistent with the result by Wang *et al.* (2003b).

The CH_3Cl factor is characterized by large signals of CH_3Cl (78% and 41.2 pptv for the observations, 95% and 25.1 pptv for the model), and no O_3 variability is explained by this factor. This factor contains significant CO variability, which can imply the biomass burning influence. However, the little or negative factor correlations with C_2H_6/C_3H_8 ratio (r = 0.01 and 0.09) and CO (r = -0.06 and -0.22), and negative correlation with latitude (r = -0.34 and -0.2) may support the hypothesis of pseudo biogenic CH_3Cl emissions from the tropics (Yoshida *et al.,* 2004) rather than biomass burning.

The NO_y/PAN factor has large signals of NO_y (44% and 120 pptv for the observations, 74% and 160 pptv for the model) and PAN (68% and 96.3 pptv for the observations, 45% and 63.4 pptv for the model). This factor is not clearly separated from the hydrocarbon factor, leading to contamination of CO in the NO_y/PAN factor and of NO_y and PAN in the hydrocarbon factor (Figure 5.2). The NO_y/PAN factor is the second important factor for tropospheric O_3 variability at mid latitudes (25.6% and 6.6 ppbv for the observations, 18.4% and 4.4 ppbv for the model). The positive factor correlations with latitude (r = 0.32 and 0.32 for the observed and simulated datasets, respectively) and CO (r = 0.46 and 0.23, respectively) and small or negative correlations with altitude (r = -0.12 and -0.1) and C_2H_6/C_3H_8 ratio (r = -0.23 and 0.07) imply that the air masses are strongly influenced by surface emissions of NO_x and other pollutants from Europe and Asia.

The hydrocarbon factor is characterized by large variability of CO (48% and 26.8 ppbv for the observations, 54% and 21. 3 ppbv for the model) and C_3H_8 (90% and 360 pptv for the observations, 90% and 272 pptv for the model). There is no contribution to tropospheric O_3 variability. It positively correlates with latitude (r = 0.52 and 0.31 for the observed and simulated datasets, respectively) and CO (r = 0.72 and 0.74, respectively). Significant negative correlations with altitudes (r = -0.65 and -0.46, respectively) and C_2H_6/C_3H_8 ratio (r = -0.64 and -0.8, respectively) reflect the influence of relatively fresh emissions from the surface (Table 5.2).

Table 5.2. The factor scores correlations (r) with latitude, altitude, C_2H_6/C_3H_8, and CO for TOPSE mid latitudes.

R	Latitude		Altitude		C_2H_6/C_3H_8		CO	
Factors	Obs	Mod	Obs	Mod	Obs	Mod	Obs	Mod
^{7}Be	-0.26	-0.37	0.42	0.62	0.02	0.1	-0.18	-0.33
θ	-0.59	-0.51	0.65	0.75	0.61	0.43	-0.64	-0.52
CH_3Cl	-0.34	-0.2	0.38	0.31	0.01	0.09	-0.06	-0.22
NO_y/PAN	0.32	0.32	-0.12	-0.1	-0.23	0.07	0.46	0.23
HC[1]	0.52	0.31	-0.65	-0.46	-0.64	-0.80	0.72	0.74

Extreme factor scores (outside 2 σ range) and the measurements that have O_3 greater than100 ppbv are excluded.
[1]HC denotes hydrocarbon factor.

5.2.1.2. TOPSE at high latitudes

Just as at mid latitudes, five factors are identified for high latitudes (^{7}Be, θ, CH_3Cl, NOy/PAN, and Hydrocarbons; Figure 5.4). The ^{7}Be factor represents the stratospheric contributions, and it shows the comparable O_3 variabilities between observations and simulation (34% and 13.7 ppbv for the observations, 45% and 12.4 ppbv for the model).

However, there is also significant difference in ^{7}Be variability between observations and simulation (415 fCi/SCM and 206 fCi/SCM, respectively), reflecting serious underestimation of ^{7}Be by GEOS-Chem. Liu *et al.,* (2001) separated stratospheric from tropospheric production of ^{7}Be in the model with locally diagnosed thermal tropopause and simulated stratospheric ^{7}Be as a separate tracer in GEOS-Chem. Based on observed ^{7}Be/^{90}Sr ratio (Dutkiewicz and Husain, 1985), 23 – 27% of ^{7}Be at surface in Northern Hemisphere is stratospheric origin. Liu *et al*., (2001) used this to constrain the model and artificially scaled down the stratospheric ^{7}Be source by a factor of 3, adjusting the excessive cross-tropopause transport in GEOS-3 until the simulated surface ^{7}Be/^{90}Sr ratio reached the observation. However, the simulated ^{7}Be mean concentrations and variabilities, and that of the ^{7}Be factor, show consistent underestimation by a factor of 2 – 3 throughout most of the dataset (to be shown), which implies that scaling down the stratospheric ^{7}Be source for fixing excessive cross-tropopause transport is unlikely to be a suitable solution for the TOPSE dataset. The stratospheric O_3 fraction from the tagged O_3 simulation suggests that ~70% is of the stratospheric origin (Figure 5.4). The tropospheric fraction is ~30% in this factor. Just as at mid latitudes, a higher O_3 variability is shown in Figure 5.4 when we include measurements with lower stratospheric O_3 (O_3 > 100). The resulting scaled EV of simulated O_3 is lower due to larger observed O_3 variability in Figure 5.5. The positive correlations with altitude (r = 0.48 and 0.4, respectively) and negative correlations with CO (r = -0.41 and -0.38) support its lower stratospheric origin (Table 5.3).

The potential temperature (θ) factor shows large variabilities of θ (77% and 29 K for the observations, 98% and 32.6 K for the model). Its contributions to O_3 levels are as

much as that of the ^{7}Be factor (39% and 15.4 ppbv for the observations, 41% and 10. 7 ppbv for the model), which is different from mid latitudes. The factor correlations in Table 5.3 may characterize this factor as long-range transport of O_3 without its precursor $NO_x$from mid to high latitudes (r = -0.1 and -0.2 with latitude, r = 0.51 and 0.55 with altitudes).

The CH_3Cl factor is characterized by large signals of CH_3Cl (78% and 38.6 pptv for the observations, 98% and 20.3 pptv for the model), but its contribution to O_3 variability is insignificant. This factor contains only significant CO variability in observations, reflecting the difference in the covariance structure of observations and simulations (Figure 5.3). According to the factor correlations (r = -0.07 and -0.2 for the observed and simulated datasets with latitude, respectively; r = 0.38 and 0.46 with altitude; r = -0.03 and -0.28 with CO), this factor is associated with long-range transport of air masses from mid to high latitudes, likely relating to the long-range influence of pseudo biogenic CH_3Cl at mid latitudes.

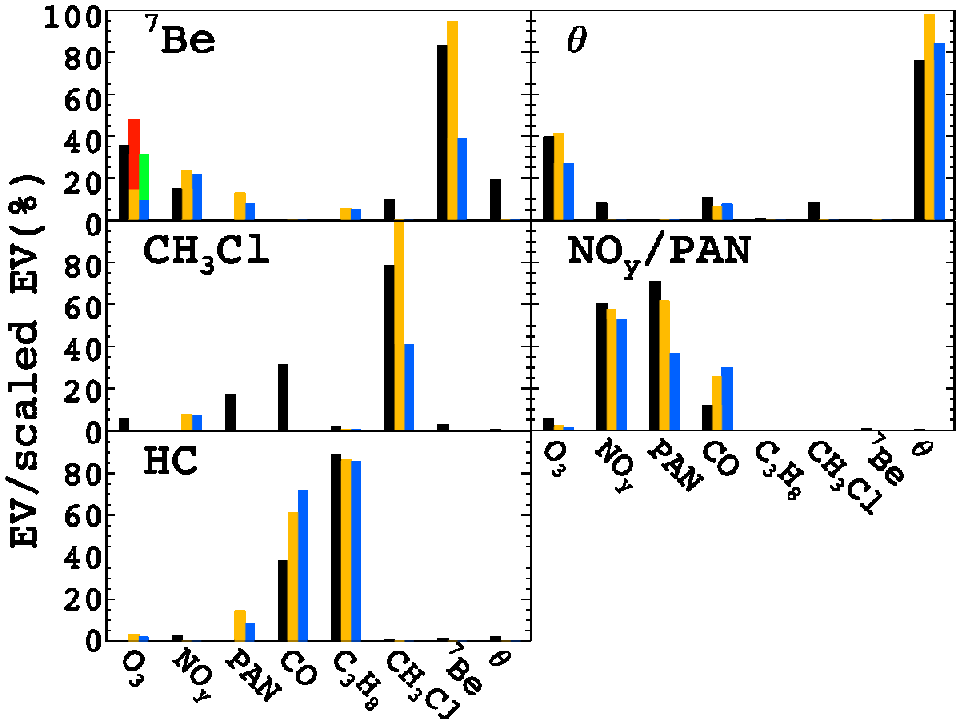


Figure 5.4. Same as Figure 5.2, but for TOPSE high latitudes.

The NO_y/PAN factor has large signals of NO_y (60% and 135 pptv for the observations, 57% and 119 pptv for the model) and PAN (72% and 132 pptv for the observations, 62% and 67.5 pptv for the model). The positive factor correlations with CO (r = 0.36 and 0.37 for the observed and simulated datasets, respectively), and the small or negative correlations with altitude (r = 0.02 and -0.34), imply that the air masses are influenced by industrial/fossil fuel emissions at high latitudes (Table 5.3). This factor, however, contributes to less than 5% of O_3 variability, reflecting less photochemical environment at high latitudes (Wang *at al*., 2003a).

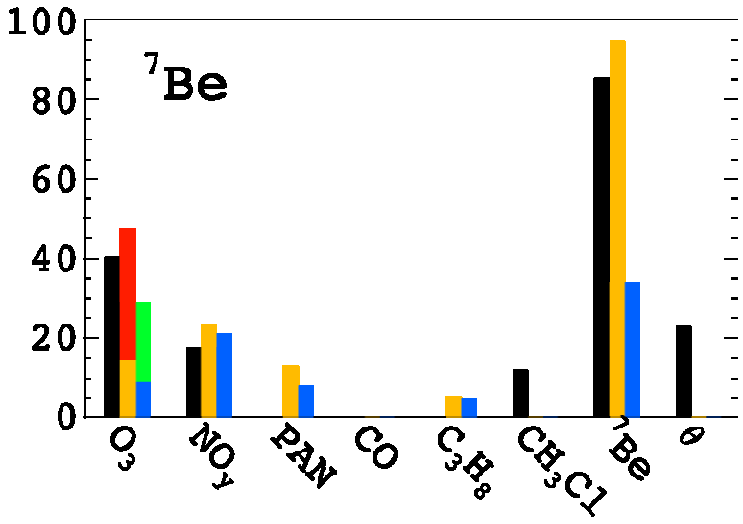


Figure 5.5. Same as Figure 5.3, but for TOPSE high latitudes (60 – 85°N).

The hydrocarbon factor is characterized by a large variability of CO (40% and 12.2 ppbv for the observations, 62% and 22.5 ppbv for the model) and C_3H_8 (90% and 314 pptv for the observations, 88% and 297 pptv for the model). Similar to mid latitudes, there is no contribution to tropospheric O_3 variability. It positively correlates with CO (r = 0.46 and 0.77, respectively). Just as mid latitudes, the significant negative correlations with altitude (r = -0.43 and -0.49, respectively) and C_2H_6/C_3H_8 ratio (r = -0.77 and -0.79, respectively) reflect air masses affected by relatively fresh emissions at high latitudes (Table 5.3).

Table 5.3. Same as Table 5.2, but for TOPSE high latitudes.

R	Latitude		Altitude		C_2H_6/C_3H_8		CO	
Factors	Obs	Mod	Obs	Mod	Obs	Mod	Obs	Mod
^7Be	-0.08	-0.08	0.48	0.4	0.25	0.11	-0.41	-0.38
θ	-0.1	-0.2	0.51	0.55	0.22	0.4	0.11	-0.62
CH_3Cl	-0.07	-0.2	0.38	0.46	0.18	-0.18	-0.03	-0.28
NO_y/PAN	0.16	0.07	0.02	-0.34	0.25	-0.14	0.36	0.37
HC	-0.08	0.04	-0.43	-0.49	-0.77	-0.79	0.46	0.77

5.3.1.3. Springtime O_3 Trends at northern mid and high latitudes

Understanding the contributions to the seasonal O_3 trend is another important purpose of this study. As stated in section 5.2.1, this study analyzed only eight tracers due to the limited availability of simulated tracers, while the previous study included fourteen tracers with seven factors (Wang *et al*., 2003b, hereafter referred to as the previous study). At mid latitudes, the seasonal increase of all factors of measurements is 6.48 ppbv/month (Table 4), consistent with the previous study (6.3 ppbv/month). The largest contributor to the O_3 seasonal trend is the NO_y/PAN factor (3.55 ppbv/month, Table 5.4) followed by the ^7Be factor (2.66 ppbv/month). That is also consistent with the previous study (3.5 ppbv/month, and 2.5 ppbv/month, respectively). In contrast, the simulated overall seasonal increase is only 3.01 ppbv/month, indicating significant underestimation. The increase from the NO_y/PAN factor is underestimated (1.32 ppbv/month), and the the ^7Be factor increase is much smaller than that of observation (1.29 ppbv/month).

At high latitudes, the overall seasonal increase from the measurements is 4.29 ppbv/month (Table 5.4), comparable with the previous study (4.6 ppbv/month). In comparison, the simulated increase is only 1.3 ppbv/month, also indicating a significant underestimation. The most contributions to the seasonal increases at high latitudes are from ^{7}Be, θ, and NO_y/PAN factors (1.78, 1.16, and 1.10 ppbv/month, respectively). However, the contributions of ^{7}Be and θ factors are significantly different from those of the previous study (0.8 ppbv/month and 0.6 ppbv/month, respectively). The previous study had additional tracers resulting in the CH_4–halocarbon factor. It accounts for transport from lower latitudes, which contributes to the largest increase of O_3 at 1.7 ppbv/month at high latitudes (Wang *et al*., 2003b). In this study, that large increase trend is mixed into the ^{7}Be and θ factors. The CH_4 – halocarbon factor contribution to O_3 variability is, however, < 10% (3 ppbv) at high latitudes; thus the effect of the missing factor on factor apportioned O_3 variability is fairly insignificant in this study. The NOy/PAN factor trend is important at high latitudes (1.10 ppbv/month) consisitent with the previous study, but the simulated seasonal increase is seriously underestimated (0.11 ppbv/month).

During TOPSE, the major contributions to the seasonal O_3 increase in springtime is from intercontinental transport of polluted air masses, while the major contributions to O_3 variability is from the stratospheric influences and long-range transport of O_3 from mid to low latitudes. These underestimations of the springtime enhancements in the model are also supported by Wang *et al.* (2006). Considering comparable concentrations and variabilites of tracers in those factors, the model probably needs to improve the schemes regarding cross-tropopause air fluxes and intercontinental transport.

Table 5.4. Factor contributions to O_3 seasonal increase (ppbv / month) for TOPSE.

	Mid latitudes		High latitudes	
	Observation	Model	Observation	Model
^7Be	2.66	1.29	1.78	0.76
θ	0.27	0.38	1.16	0.77
CH_3Cl	0	0	0.27	0
NO_y/PAN	3.55	1.32	1.10	0.11
HC	0	-0.01	-0.02	-0.34
Total	6.48	3.01	4.29	1.30

Only the measurements O_3 < 100 ppbv are analyzed.

5.3.2. TRACE-P

The TRACE-P experiment was conducted in order to investigate the effects of Asian outflow to the Pacific during spring (Jacob *et al*., 2003). As mentioned in section 5.2.1, the TRACE-P results are biased toward the middle and upper troposphere (more than 40% of the data is above 7 km). The processes of the analysis are same as TOPSE. The PMF result in TRACE-P is more sensitive and is likely biased by outliers in the dataset, since the available number of data with coincident ^7Be measurements is only 65 and 78 at mid and low latitudes, respectively.

5.3.2.1. TRACE-P at mid latitudes

Four factors are identified for mid latitudes (^7Be, θ, CH_3Cl, and NOy/Hydrocarbons, Figure 5.6). The ^7Be factor shows relatively larger O_3 variability in the observations than model results (68% and 20.8 ppbv for the observations, 48.5 % and 13.8 ppbv for the model). There is also a significant underestimation in simulated ^7Be variability (428 fCi/SCM and 211 fCi/SCM, respectively) (also see Table 5.1 for mean values). The

stratospheric O_3 fraction from the tagged O_3 simulation represents ~80% of the stratospheric origin. As in Figure 5.2, much larger observed O_3 variability is shown in Figure 5.7 when we include measurements with O_3 above 100 ppbv. Much larger observed ^{7}Be variability is also shown in Figure 5.7 (1150 and 270 fCi/SCM for the observed and simulated datasets, respectively) for the same reason. Positive correlation with altitudes (r = 0.12 and 0.63, respectively) and small and negative correlation with CO (r = -0.2 and -0.66) support its lower stratospheric origin (Table 5.5). The notable difference between observation and simulation is that most of O_3 variability is solely explained by ^{7}Be factor in the observations (68%) with a large unidentified fraction (31.7%), while the simulation shows the factor contributions to O_3 variability from ^{7}Be, θ, and NO_y/PAN factors (48.5 %, 25.6%, and 14.5 %, respectively). As mentioned in section 5.3.1, this discrepancy represents the problems of the limited number of data due to ^{7}Be availability. The consistent result is that the stratospheric influence on O_3 variabilities of both the observations and the simulation is largest in TRACE-P mid latitudes (30 – 45 °N), supporting the hypothesis of active interchange of air masses between stratosphere and troposphere at mid latitudes (Holton *et al*., 1995; Chen, 1995).

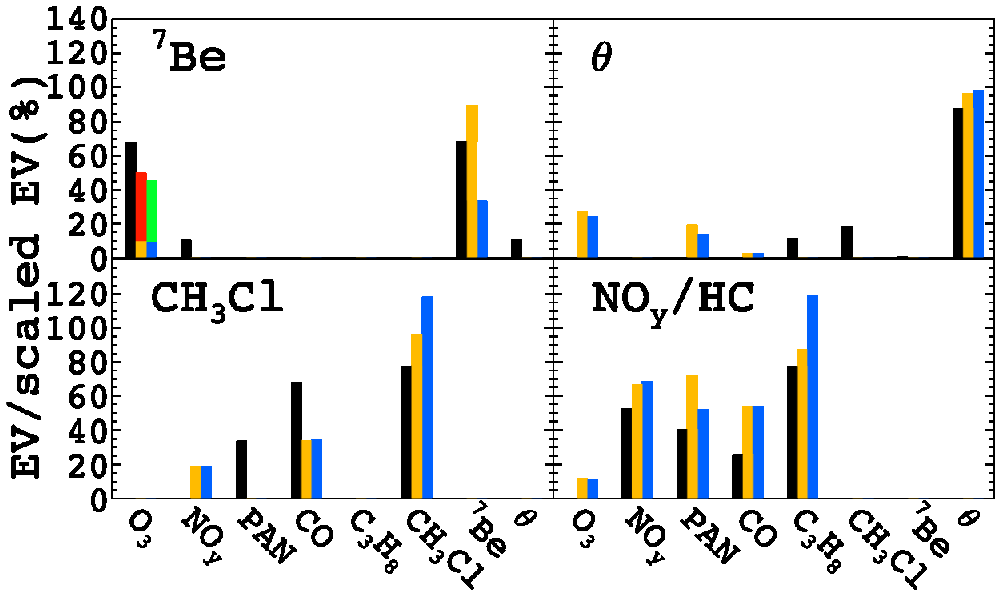


Figure 5.6. Same as Figure 5.2, but for TRACE-P mid latitudes (30 – 45°N).

The potential temperature (θ) factor shows large signals of θ (87% and 26.3 K for the observations, 93% and 23.4 K for the model). The factor correlation in Table 5.5 may characterize this factor as long-range transport of air masses from the tropics (r = -0.33 and -0.67 with latitude; r = 0.79 and 0.67 with altitude; and r = 0.86 and 0.82 with C_2H_6/C_3H_8 ratio).

The CH_3Cl factor is characterized by the large signals of CH_3Cl (76% and 37.4 pptv for the observations, 95% and 42.2 pptv for the model), and the contribution to O_3 variability is insignificant by this factor. This factor also contains significant CO variability (65% and 74.5 ppbv for the observations, 38% and 26.6 ppbv for the model), NO_y vaiability (168 pptv only in simulation), and PAN variability (164 pptv only in observation). The discrepancies of NO_y and PAN variability, both in the observations and

simulation, reflect their differing covariance structures (Figure 5.6). According to the factor correlations (r = -0.25 and -0.77 with altitude; r = -0.33 and -0.6 with C_2H_6/C_3H_8 ratio; and r = 0.85 and 0.88 with CO), this factor is likely associated with relatively fresh biomass burning emissions from the surface (Table 5.5), which is different from that of TOPSE (section 5.3.1).

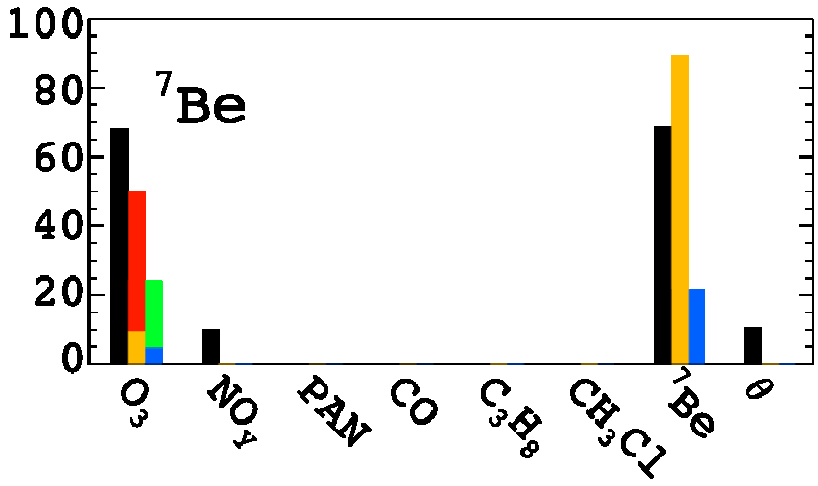


Figure 5.7. Same as Figure 5.3, but for TRACE-P mid latitudes.

In TRACE-P analysis, the NO_y/PAN and hydrocarbon factors are combined (now NO_y/hydrocarbon factor) because the separation of those factors leads to incomparable factor profiles between the measurements and model results. The NO_y/hydrocarbon factor is characterized by a large variability of NO_y (51% and 509 pptv for the observations, 67% and 523 pptv for the model) PAN (40% and 195 pptv for the observations, 72% and 234 pptv for the model), CO (25% and 27.8 ppbv for the observations, 53% and 54.6 ppbv for the model), and C_3H_8 (78% and 335 pptv for the observations, 87% and 461

pptv for the model). This factor shows a contribution to tropospheric O_3 variability only in the simulation (14.5%) and highly positive factor correlations with CO (r = 0.85 and 0.8, respectively). Significant negative correlations with altitudes (r = -0.77 and -0.31, respectively) and the C_2H_6/C_3H_8 ratio (r = -0.64 and -0.16, respectively) may reflect relatively fresh industrial/fossil fuel emissions over Asia (Table 5.5).

Table 5.5. Same as Table 5.2, but for TRACE-P mid latitudes.

R	Latitude		Altitude		C_2H_6/C_3H_8		CO	
Factors	Obs	Mod	Obs	Mod	Obs	Mod	Obs	Mod
^7Be	-0.37	0.05	0.12	0.63	0.32	0.28	-0.2	-0.65
θ	-0.33	-0.67	0.79	0.67	0.86	0.82	-0.67	-0.36
CH_3Cl	-0.02	0.11	-0.25	-0.77	-0.33	-0.6	0.85	0.88
NO_y/PAN	0.18	-0.42	-0.77	-0.31	-0.64	-0.16	0.85	0.8

5.3.2.2. TRACE-P at low latitudes

Four factors also are identified for low latitudes (^7Be, θ, CH_3Cl, and NO_y/hydrocarbons, Figure 5.8). The ^7Be factor shows smaller O_3 variability in the measurements than the model simulation (17.4% and 7.1 ppbv for the observations, and 30.8 % and 9.9 ppbv for the model). Significant underestimation is found in simulated ^7Be variability (411 fCi/SCM and 139 fCi/SCM for the observed and simulated datasets, respectively). The stratospheric O_3 fraction from the tagged O_3 simulation represents ~50% of the stratospheric origin, the smallest stratospheric influence among the datasets. There are no data with O_3 above 100 ppbv in both observations and simulation at low

latitudes. The positive factor correlations with altitude (r = 0.39 and 0.67, respectively) and negative correlations with CO (r = -0.12 and -0.59) reflect its lower stratospheric origin (Table 5.6).

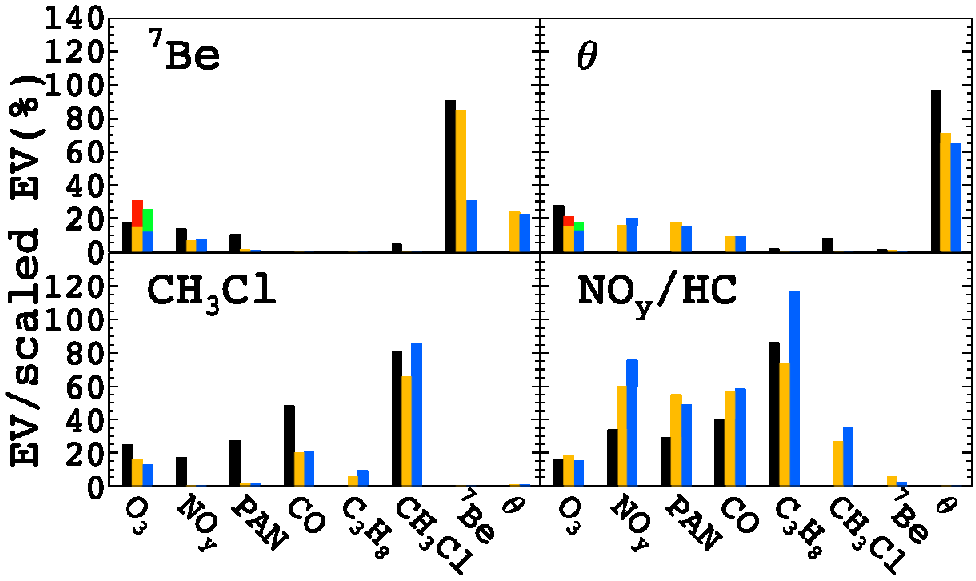


Figure 5.8. Same as Figure 5.2, but for TRACE-P low latitudes.

The potential temperature (θ) factor shows large signals of θ (96% and 34.5 K for the observations, 77% and 23.2 K for the model). This factor contains small signals of simulated NO_y, PAN, and CO, representing different covariance structure from the observations. The factor correlation may characterize this factor as long-range transport of air masses from the tropics in the free troposphere (r = -0.05 and 0 with latitude, r = 0.98 and 0.98 with altitude, and r = 0.61 and 0.52 with C_2H_6/C_3H_8 ratio). While the ^{7}Be factor is the largest contributor to simulated O_3 variability at low latitudes, the θ factor is the largest contributor to observed O_3 variability (27.4% and 11.1 ppbv for the

observations 21% and 6.7 ppbv for the model). ~20% of stratospheric fraction in O_3 variability (Figure 5.8) represents the contamination in this factor.

The CH_3Cl factor is characterized by large signals of CH_3Cl (81% and 39.7 pptv for the observations, 67% and 42 pptv for the model) and a significant contribution to O_3 variability is found in this factor (24.6% and 9.9 ppbv for the observations, 15.7% and 5 ppbv for the model). Similar to the θ factor, the CH_3Cl factor shows relatively more contributions to O_3 in the observations than the model. Only a simulated CH_3Cl signal is found in NO_y/hydrocarbon factors, and the CH_3Cl factor also contains significant CO (50% and 40.7 ppbv for the observaitons, 22% and 17.4 ppbv for the model), NO_y (53.7 pptv), and PAN (52 pptv) variabilities only in observations due to differences in covariance structure. According to the factor correlations (r = 0.22 and 0.4 with latitude; r = 0.03 and 0.13 with altitude; r = -0.33 and -0.4 with C_2H_6/C_3H_8 ratio; and r = 0.53 and 0.1 with CO), this factor reflects relatively fresh biomass burning emissions from the subtropical regions (Table 5.6), similar to the TRACE-P mid latitudes.

Table 5.6. Same as Table 5.2, but for TRACE-P low latitudes.

R	Latitude		Altitude		C_2H_6/C_3H_8		CO	
Factors	Obs	Mod	Obs	Mod	Obs	Mod	Obs	Mod
^7Be	0.33	-0.06	0.39	0.67	0.06	0.41	-0.12	-0.59
θ	-0.05	0	0.98	0.93	0.61	0.52	-0.66	-0.49
CH_3Cl	0.22	0.4	0.03	0.13	-0.33	-0.4	0.53	0.1
NO_y/PAN	0.25	0.35	-0.65	-0.53	-0.68	-0.75	0.64	0.95

NO_y/hydrocarbon factor is characterized by large variabilities of NO_y (33% and 185 pptv for the observations, 60% and 376 pptv for the model), PAN (30% and 56.5 pptv for the observations, 55% and 94.3 pptv for the model), CO (40% and 34.2 ppbv for the observations, 57% and 49.6 ppbv for the model), and C_3H_8 (85% and 172 pptv for the observations, 75% and 234 pptv for the model). This factor also shows significant contributions to tropospheric O_3 variability (16% and 6.5 ppbv for the observations, 18.5% and 6 ppbv for the model) and positive factor correlations with latitude (r = 0.25 and 0.35, respectively) and CO (r = 0.64 and 0.95, respectively), and significant negative correlations with altitude (r = -0.65 and -0.53, respectively) and C_2H_6/C_3H_8 ratio (r = -0.68 and -0.75, respectively) indicating fresh polluted air plumes from East Asia (Table 5.6).

5.4. Conclusions.

Trace gas measurements of TOPSE and TRACE-P experiments are analyzed with PMF, an advanced factor analysis, in order to evaluate model simulated source contributions to O_3 variability and the springtime increase. We select a suite of relatively long-lived variables that includes seven chemicals (O_3, NO_y, PAN, CO, C_3H_8, CH_3Cl, and ^7Be) and one dynamic tracer (potential temperature). The evaluation has a bias towards a high altitude of 5 – 8 km (~70 % of the data) for TOPSE and 7 – 12 km (~50 % of the data) for TRACE-P, due to the availability of ^7Be measurements. Since the emissions and photochemical environment are quite different with latitude, we divided these two aircraft experiment regions by latitude: mid and high latitudes in TOPSE (40 – 60 °N and 60 – 85 °N, respectively), and mid and low latitudes in TRACE-P (30 – 45 °N

and 15 – 30 °N, respectively). This study based on regional aircraft measurements could represent the global source contributions to tropospheric O_3 depending on latitudes. However, the significant local variations in stratospheric contributions and emissions of O_3 precursors give large uncertainties in the global implications of tropospheric O_3 sources from this study. In general, the mean concentrations of simulated tracers are comparable with those of observed tracers. However, the significant underestimation of ^{7}Be is found in the most dataset by a factor of 2 – 3, implying the problem in scaling down of stratospheric ^{7}Be sources by a factor of ~3 to correct excessive cross-tropopause transport (Liu *et al*., 2001).

Five factors are resolved in TOPSE (^{7}Be, θ, CH_3Cl, NO_y/PAN, and Hydrocarbons). We investigate the factor correlations with altitude, latitude, C_2H_6/C_3H_8 ratio, and CO to help identify the characteristics of each factor. In general, the factor profiles of observations and simulations are fairly comparable, suggesting that the model captures major source factors for tropospheric O_3. At mid latitudes, the ^{7}Be factor shows large stratospheric contributions to tropospheric O_3 variability (40% and 20 ppbv for the observations, 58% and 14 ppbv for the model). The θ factor represents intercontinental long-range transport of O_3 from low to mid latitudes, contributing 14% of observed O_3 variability (3.6 ppbv) and 12.8% of simulated O_3 variability (3.1 ppbv). The NO_y/PAN factor is an important contributor to tropospheric O_3 variability at mid latitudes (25.6% and 6.6 ppbv for the observations, 18.4% and 4.4 ppbv for the model), characterizing the contributions from surface NOx emissions from Europe and Asia.

At high latitudes, the ^{7}Be factor shows a significant contribution to O_3 variability (34% and 13.7 ppbv for the observations, 45% and 12.4 ppbv for the model) that is

comparable to that of mid latitudes. However, the contribution of the θ factor to O_3 variability at high latitudes is significantly larger than at mid latitudes, which implies that intercontinental transport of O_3 from mid to high latitudes is of more importance for O_3 variability at high latitudes (39% and 15.4 ppbv for the observations, 41% and 10.7 ppbv for the model). The CH_3Cl factor at high latitudes likely reflects the long-range influence of biogenic CH_3Cl emissions from lower latitudes, with 5.4% contributions to only observed O_3 variability. The NO_y/PAN factor contributes 5.6% of the observed O_3 variability at high latitudes.

At mid latitudes, the average O_3 variations are fairly comparable between observation and simulation (25.7 ppbv and 24 ppbv for the observed and simulated datasets, respectively), but there is more contribution from NO_y/PAN factor (6.6 ppbv and 4.4 ppbv, respectively) and less from the ^{7}Be factor (10.2 ppbv and 14 ppbv, respectively) in the observations (Figure 5.6). Compared with mid latitudes, the average O_3 variations are underestimated at high latitudes (38 ppbv and 26.9 ppbv, respectively). Those discrepancies are mostly due to the underestimations of θ, CH_3Cl, and NO_y/PAN factor contributions (Figure 5.9), apparently representing model-underestimation of springtime O_3 contributions from intercontinental transport of O_3 and its precursors.

A substantial underestimation by the model is shown in the smaller seasonal increasing trends (observed 6.48 ppbv/month, and simulated 3.0 ppbv/month for mid latitudes, and observed 4.3 ppbv/month and simulated 1.3 ppbv/month for high latitudes) during TOPSE. At mid latitudes, the observed seasonal increase is driven mostly by the NO_y/PAN factor (3.55 ppbv/month) and the ^{7}Be factor (2.66 ppbv/month), suggesting that intercontinental transport of polluted air that contributes to the seasonal increase

more than stratospheric influence does. Simulated O_3 seasonal increasing trend at mid latitudes also is attributed mostly to the NO_y/PAN and the ^7Be factors, but their magnitudes are significantly underestimated (1.32 ppbv/month and 1.29 ppbv/month, respectively). At high latitudes, the observed seasonal O_3 increase is attributed to the ^7Be (1.78 ppbv/month), θ (1.16 ppbv/month), and NOy/PAN factors (1.10 ppbv/month). The seasonal O_3 increase is underestimated by the model, and that of NO_y/PAN is much smaller (0.11 ppbv/month). These discrepancies apparently indicate that the GEOS-Chem simulation cannot reproduce the dramatic seasonal O_3 increase in the free troposphere at northern mid and high latitudes.

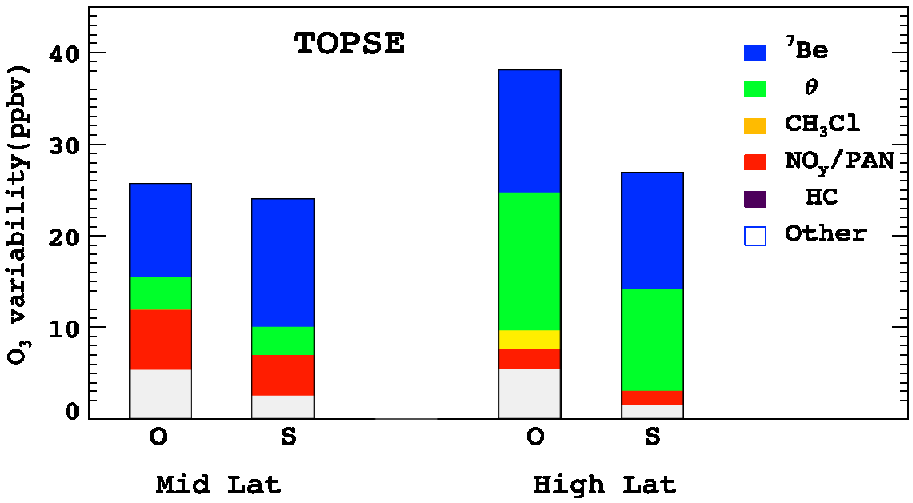


Figure 5.9. O_3 variation ($[O_3]_{average}$ - $[O_3]_{min}$) by factors for TOPSE. Left two bars are for mid latitudes and right two bars are for high latitudes. "O" and "S" denote "observation" and "simulation" respectively. The each factor contribution is expressed by different color. "Other" denotes the unexplained fractions by PMF.

For TRACE-P, four factors are resolved (^{7}Be, θ, CH_3Cl, and NO_y/hydrocarbons). At mid latitudes, the ^{7}Be factor represents the stratospheric O_3 influence and is the largest contributor (68% and 20.8 ppbv for the observations, 48.5% and 13.8 ppbv for the model). The observation only shows substantial stratospheric contributions (68%, 20.8 ppbv), while the model results include θ and NO_y/hydrocarbon factor contributions (25.6% and 7.3 ppbv for the observations, 14.5% and 4.1 ppbv for the model).

At low latitudes, the ^{7}Be factor contributions (stratospheric influence) are smaller than other data sets (17.4% and 7.1 ppbv for the observations, 30.8% and 9.9 ppbv for the model). The θ factor is characterized by long-range transport through the tropical free troposphere, and its contribution to O_3 variability is large (27.4% and 11.1 ppbv for the observations, 21% and 6.7 ppbv for the model). The characteristic of the CH_3Cl factor in TRACE-P is likely related to biomass burning due to the positive factor correlation with CO; it accounts for 24.6% (9.9 ppbv) in observed and 15.7% (5 ppbv) in simulated O_3 variability. The NO_y/hydrocarbon factor (industry/urban emissions) contributes to O_3 variability (16%, 6.5 ppbv, and 18.5%, 6 ppbv, respectively). The overall tropospheric contribution to O_3 variability is largest (68% and 55.2%, respectively) at TRACE-P low latitudes.

At mid latitudes, the average O_3 variations are fairly comparable (30.6 ppbv and 28.3 ppbv, respectively), with the largest fractions for ^{7}Be factors (20.8 ppbv and 13.8 ppbv, respectively) (Figure 5.10). The observed average O_3 variation is larger than that of simulation, due mostly to the underestimation of contributions by long-range transport of O_3 and biomass burning. The stratospheric O_3 influence is smallest at low latitudes and the model overestimates this contribution by 2.8 ppbv.

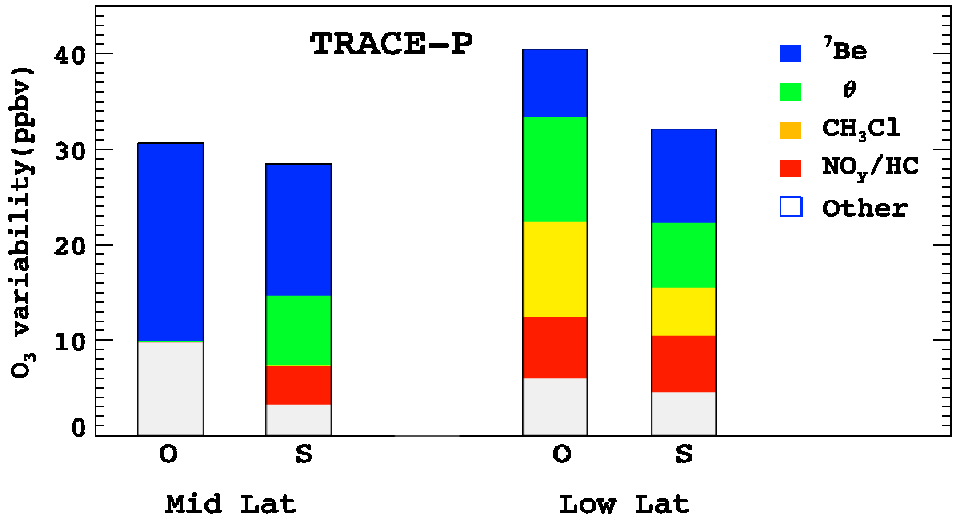


Figure 5.10. Same as Figure 5.9, but for TRACE-P. Left two bars are for mid latitudes and right two bars for low latitudes.

There are a number of limitations in this work. First, the suitable set of data only has eight tracers due to the limited number of simulated tracers, and the amount of coincidental data is rather small due to the availability of ^{7}Be measurements. Second, the relatively coarse spatial resolution of the model might contribute to the uncertainties in the results. The major problems identified in the model include the tendency to underestimate the contribution to O_3 and weaker seasonal trend by intercontinental transport of pollutant and spring increase of O_3 from the stratosphere even though the overall magnitude appears to be reasonable.

REFERENCES

Abbot, D.S., P.I. Palmer, R.V. Martin, K.V. Chance, D.J. Jacob, and A. Guenther (2003). "Seasonal and interannual variability of North American isoprene emissions as determined by formaldehyde column measurements from space." *Geophysical Reserch Letter* 30(17): 1886, doi:10.1029/2003GL017336.

Allen, D. J., J. E. Dibb, B. Ridley, K. E. Pickering, and R. W. Talbot (2003). "An estimate of the stratospheric contribution to springtime tropospheric ozone maxima using TOPSE measurments and beryllium-7 simulations." *Journal of Geophysical Research* 108: (D4), 8355, doi: 10.1029/2001JD001428.

Andreae, M.O., and P. Merlet (2001). "Emission of trace gases and aerosols from biomass burning." *Global Biogeochemical Cycles* 15: 955-966.

Atkinson, R. (1994). "Gas-phase tropospheric chemistry of organic compounds." *Journal of Physical Chemistry Reference Data Monograph* 2: 13-46.

Atlas, E. L., B. A. Ridley, and C. Cantrell (2003). "Tropospheric Ozone Production about the Spring Equinox (TOPSE) Experiment: Introduction." *Journal of Geophysical Research* 108: (D4), 8353, doi:10.1029/2002JD003172.

Berntsen, T. K., S. Karlsdo´ttir, and D. A. Jaffe (1999). "Influence of Asian emissions on the composition of air reaching the northwestern United States." *Geophysical Research Letter* 26: 2171–2174.

Bey, I., D.J. Jacob, R.M. Yantosca, J.A. Logan, B.D. Field, A.M. Fiore, Q. Li, H.Y. Liu, L.J. Mickley, and M.G. Schultz (2001). "Global modeling of troposheric chemistry with assimilated meteorology: Model description and evaluation." *Journal of Geophysical Research* 106: 23,073-23,096.

Bey, I., D. J. Jacob, J. A. Logan, and R. M. Yantosca (2001). "Asian chemical outflow to the Pacific in spring: Origins, pathways and budgets." *Journal of Geophysical Research* 106: 23,097– 23,114.

Blake, N. J., D. R. Blake, B. C. Sive, A. S. Katzenstein, S. Meinardi, O. W. Wingenter, E. L. Atlas, F. Flocke, B. A. Ridley, and F. S. Rowland (2003). "The seasonal evolution of NMHCs and light alkyl nitrates at middle to high northern latitudes during TOPSE." *Journal of Geophysical Research* 108: (D4), 8359, doi:10.1029/2001JD001467.

Brasseur G. P., J. J. Orlando, and G. S. Tyndall (1999). "Atmospheric Chemistry and Global Change" *Oxford University Press, Inc*., New York, New York.

Browell, E. V., *et al*. (2003). "Ozone, aerosol, potential vorticity, and trace gas trends observed at high latitudes over North America from February to May 2000." *Journal of Geophysical Research* 108: (D4), 8369, doi:10.1029/2001JD001390.

Burrows, J.P., *et al* (1999). "The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results." *Journal of Atmospheric Science* 56: 151-175.

Cantrell, C. A., J. A. Davidson, A.H. McDaniel, R.E. Shetter, and J.G. Calvert (1990). "Temperature-dependent formaldehyde cross sections in the near-ultraviolet spectral region." *Journal of Physical Chemistry* 94, 3902-3908.

Cantrell, C. A., *et al*. (2003). "Steady state free radical budgets and ozone photochemistry during TOPSE." *Journal of Geophysical Research* 108: (D4), 8361, doi:10.1029/2002JD002198.

Carpenter, L. J., T. J. Green, G. P. Mills, S. Bauguitte, S. A. Penkett, P. Zanis, E. Schuepbach, N. Schmidbauer, P. S.Monks, and C. Zellweger (2000). "Oxidized nitrogen and ozone production efficiencies in the springtime free troposphere over the Alps." *Journal of Geophysical Research* 105: 14,547–14,559.

Carslaw, N., N. Bell, A.C. Lewis, J.B. McQuaid, and M.J. Pilling (2000). "A detailed case study of isoprene chemistry during the EASE96 Mace Head campaign." *Atmospheric Environment* 34: 2827-2836.

Chameides, W. L., F. Fehsenfeld, M. O. Rodgers, C. Cardelino, J. Martinez, D. Parrish, W. Lonneman, D. R. Lawson, R. A. Rasmussen, P. Zimmerman, J. Greenberg, P. Middleton and T. Wang (1992). "Ozone precursor relationships in the ambient atmosphere." *Journal of Geophysical Research* 97: 6037-6055.

Chance, K., P.I. Palmer, R.J.D. Spurr, R.V. Martin, T.P. Kurosu, and D.J. Jacob (2000). "Satellite observations of formaldehyde over North America from GOME." *Geosphysical Research Letter* 27: 3461-3464.

Chen, P. (1995). "Isentropic cross-tropopause mass exchange in the extratropics." *Journal of Geophysical Research* 100, 16,661 – 16,673.

Chueinta, Wanna, P.K. Hopke, Pentti Paatero (2000). "Investigation of sources of atmospheric aerosol at urban and suburban residential areas in Thailand by positive matrix factorization." *Atmospheric Envronment* 34: 3319-3329

Claeys M., B. Graham, G. Vas, W. Wang, R. Vermeylen, V. Pashynska, J. Cafmeyer, P. Guyon, M.O. Andreae, P. Artaxo, W. Maenhaut (2004). "Formation of Secondary Organic Aerosols Through Photooxidation of Isoprene." *Sience* 303: 1173-1176.

Conn E E. (1991). "The metabolism of a natural product –lessons learned from cyanogenic glycosides." *Planata Medicda, 57*, s1-s9.

de Gouw, J. A. *et al*. (1999). "Emissions of volatile organic compounds from cut grass and clover are enhanced during the drying process." *Geophysical Research Letters 26*, 811–814.

de Gouw, J. A., C. Warneke, D. D. Parrish, J. S. Holloway, M. Trainer, and F. C. Fehsenfeld (2003). "Emission source and ocean uptake of acetonitrile (CH3CN) in the atmosphere." *Journal of Geosphysical Reserach 108(D11)*, 4329, doi: 10.1029/2002JD002897, 2003.

Dibb, J. E., L. D. Meeker, R. C. Finkel, J. R. Southon, M. W. Caffee, and L. A. Barrie (1994). "Estimation of stratospheric input to the Arctic troposphere: ^{7}Be and ^{10}Be in aerosols at Alert, Canada." *Journal of Geophysical Research* 99: 12,855–12,864.

Dibb, J. E., R. W. Talbot, E. Scheuer, G. Seid, L. DeBell, B. Lefer, and B. Ridley (2003). "Stratospheric influence on the northern North American free troposphere during TOPSE: 7Be as a stratospheric tracer." *Journal of Geophysical Research* 108(D4), 8363,doi:10.1029/2001JD001347.

Draxler, R. R., and G. D. Hess (1998). "An overview of the HY-SPLIT 4 Modeling System for trajectories, dispersion, and deposition."*Australian Meteorological Magazine* 47: 295– 308.

Dumdei, B. E., Kenny, D. V., Shepson, P. B., Kleindienst, T. E., Nero, C. M., Cupitt, L. T., and Claxton, L. D. (1988). "Analysis of the Products of Toluene Photooxidation and Measurement of Their Mutagenic Activity." *Environmental Science and Technology* 22: 12, 1493–1498.

Duncan, B. N., R.V. Martin, A.C. Staudt, R. Yevich, and J.A. Logan (2003). "Interannual and seasonal variability of biomass burning emissions constrained by satellite observations." *Journal of Geosphysical Research* 108: doi: 10.1029/2002JD002378.

Dutkiewicz V.A. and L. Husain (1985). "Stratospheric and tropospheric components of ^{7}Be in surface air." *Journal of Geosphysical Research* 90: 5785 – 5788.

Enting, I (2002). "Inverse Problems in Atmospheric Constituent Transport." *Cambridge Univ. Press*, Cambridge, U.K.

Fall, R. (2003), "Abundant oxygenates in the atmosphere: A biochemical perspective." *Chemical Review 103,* 4,941–4, 951.

Fiore, A. M., D. J. Jacob, I. Bey, R. M. Yantosca, B. D. Field, and J. Wilkinson (2002). "Background ozone over the United States in summer: Origin and contribution to pollution episodes." *Journal of Geophysical Research* 107, 4279, doi:10.1029/2001JD000982.

Fuentes, J. D. *et al*. (2000). "Biogenic hydrocarbons in the atmospheric boundary layer: A review." *Bulletin of American Meteorological Society, 81,* 1,537–1,575.

Fuelberg, H. E., R. E. Newell, D. J. Westberg, J. C. Maloney, J. R. Hannan, B. D. Martin, M. A. Avery, and Y. Zhu (2001). "A meteorological overview of the second Pacific exploratory in the tropics." *Journal of Geophysical Research 106 (D23),* 32427-32443.

Fuelberg, H. E., C. M. Kiley, J. R. Hannan, D. J. Westberg, M. A. Avery, and R. E. Newell (2003). "Meteorological conditions and transport pathways during the Transport and Chemical Evolution over the Pacific (TRACE-P) experiment." *Journal of Geophysical Research 108(D20),* 8782, doi:10.1029/2002JD003092.

Fuhrer, J. (2002). "Ozone impacts on vegetation." *Ozone Science and Engineering* 24: 69-74.

Fukui, Y., and P. V. Doskey (1998). "Air-surface exchange of nonmethane organic compounds at a grassland site: Seasonal variations and stressed emissions." *Journal of Geophysical Research* 103, 13,153–13,168.

Galbally, I. E., and W. Kirstine (2002). "The production of methanol by flowering plants and the global cycle of methanol." *Journal of Atmospheric Chemistry 43,* 195–229.

Gleadow and Woodrow (2000). "Temporal and spatial variation in cyanogenic glycosides in Eucalyptus cladocalyx." *Tree Physiology, 20,* 591-598.

Guenther, A., C.N. Hewitt, D. Erickson, R. Fall, C. Geron, T. Graedel, P. Harley, L. Klinger, M. Lerdau, W.A. Mckay, T. Pierce, B. Scholes, R. Steinbrecher, R. Tallamraju, J. Taylor, and P. Zimmerman (1995). "A global model of natural volatile organic compound emissions." *Journal of Geosphysical Research* 100: 8873-8892.

Guenther, A. *et al*. (2000). "Natural emissions of non-methane volatile organic compounds; carbon monoxide, and oxides of nitrogen from North America." *Atmospheric Environment 34,* 2,205–2,230.

Harry, D. E., and T. W. Kimmerer (1991). "Molecular genetics and physiology of alcohol dehydrogenase in woody plants." *Forest Ecology Management 43,* 251–272.

Hatakeyama, S., K. Izumi, T. Fukuyama, H. Akimoto, and N. Washida (1991). "Reactions of OH with α-pinene and β-pinene in air: Estimate of global CO production and atmospheric oxidation of terpenes." *Journal of Geophysical Research* 96: 947-958.

Heald, C.L., *et al* (2004). "Comparative inverse analysis of satellite (MOPITT) and aircraft (TRACE-P) observations to emstimate Asian sources of carbon monoxide." *Journal of Geosphysical Research* 109: (D23), D23306, doi:10.1029/ 2004JD00515.

Heikes, B. G. *et al*. (2002). "Atmospheric methanol budget and ocean implication." *Global Biochemical Cycles* 16(4), 1133, doi:10.1029/2002GB001895.

Heirtzler, J. R (2002). "The future of the South Atlantic anomaly and implications for radiation damage in space." *Journal of Atmospheric Solar-Terrerstrial Physics* 64: 1701–1708.

Helmig, D., *et al* (1998). "Vertical profiling and determination of landscape fluxes of biogenic nonmethane hydrocarbons within the planetary boundary layer in the Peruvian Amazon." *Journal of Geosphysical Research* 103: 25,519-25,432.

Herman, J. R., and E. A. Celarier (1997). "Earth surface reflectivity climatology at 340–380 nm from TOMS data." *Journal of Geophysical Research* 102, 28,003–28,011.

Hewitt, C.N., and R. Street (1992). "A qualitative assessment of the emission of non-methane hydrocarbons from the biosphere to the atmosphere in the U.K.: Present knowledge and uncertainties." *Atmospheric Evironment* 26: 3067-3077.

Holton, J. R., P. H. Haynes, E. M. McInyre, A. R. Douglass, R. B. Rood, and L. Pfister (1995). "Stratosphere-troposphere exchange" *Review of Geophysics* 33: 403-439.

Hopke, P. K. (1985). "Receptor modeling in environmental chemistry." Wiley, New York.

Horowitz, L. W., J. Liang, G. M. Gardner, and D.J. Jacob (1998). "Export of reactive nitrogen from North America during summertime: Sensitivity to hydrocarbon chemistry." *Journal of Geophysical Research* 103: 13,451-13,476.

Houweling, S., F. Dentener, and J. Lelieveld (1998). "The impact of non-methane hydrocarbon compounds on tropospheric photochemistry." *Journal of Geosphysical Research* 103: 10,673-10,696.

IPCC (2001). (Intergovernmental Panel on Climate Change). "Climate change 2001: The Scientific Basis." Cambridge, *Cambridge University Press*.

Jacob, D. J., J. A. Logan, and P. P. Murti (1999). "Effect of rising Asian emissions on surface ozone in the United States." *Geophysical Research Letter* 26: 2175– 2178.

Jacob, D. J. *et al*. (2003). "Transport and Chemical Evolution over the Pacific (TRACE-P) aircraft mission: Design, execution, and first results." *Journal of Geophysical Research 108* (D13), 1 –19.

Jacob, D. J. *et al*. (2002). "Atmospheric budget of acetone." *Journal of Geophysical Research 107* (D12), 10.1029/2001JD000694.

Jacob, D. J. *et al*. (2005). "Global budget of methanol: Constraints from atmospheric observations." *Journal of Geophysical Research 110*, D08303, doi:10.1029/2004JD005172

Jacobson, M. Z., and R. P. Turco (1994). "SMVGEAR: A sparse-matrix, vectorized Gear code for atmospheric models." *Atmospheric Environment*. 28, 273– 284.

Jaffe, D., *et al*. (1999). "Transport of Asian air pollution to North America." *Geophysical Research Letter* 26: 711 –714.

Jost, C. *et al*. (2003). "Trace gas chemistry in a young biomass burning plume over Namibia: Observations and model simulations." *Journal of Geosphysical Research 108*(D13), 8482, doi:10.1029/2002JD002431.

Kaiser, H. P. (1958). "The varimax criterion for analytical rotation in factor analysis." *Psychometrika* 23, 187– 196.

Karl, T. *et al*. (2002). "Transient releases of acetaldehyde from tree leaves – products of a pyruvate overflow mechanism?" *Plant Cell Environment 25*, 1,121–1,131.

Kamens, R. M., M.W. Gery, H. E. Jeffries, M. Jackson, and E. I. Cole (1982). "Ozone-isoprene reactions: Product formation and aerosol potential." *International Journal of Chemical Kinetics* 14: 955-975.

Kanamitsu, M. (1989). "Description of the NMC Global Data Assimilation and Forecast System." *Weather Forecasting* 4: 335– 342.

Klinger, L.F., J. Greenberg, A. Guenther, G. Tyndall, P. Zimmerman, M. M' Bangui, J.M. Moutsambot, and D. Kenfck (1998). "Patterns in volatile organic compound emissions along a savanna rainforesst gradient in central Africa." *Journal of Geophysical Reserach* 103: 1443-1454.

Koren, H. S. (1995). "Associations between criteria air pollutants and asthma." *Environmental Health Perspectives* 103: 235-242.

Kurosu, T.P., K. Chance, and R.J.D. Spurr (1999). "GRAG: Cloud Retrieval Algorithm for the European Space Agency's Global Ozone Monitoring Experiment*." in Proceedings of the European Symposium of Atmospheric Measurements From Space*, opp. 513-521, Eur. Space Agency, Paris.

Lal, D., and B. Peters (1967). "Cosmic ray produced radioactivity on the Earth." in *Handbuch der Physik,* 46/2, 551-612, edited by K. Sitte, Springer-Verlag, New York, 1967.

Lamb, B., A. Guenther, D. Gay, and H. Westberg (1987). "A national inventory of biogenic hydrocarbon emissions." *Atmospheric Environment* 21: 1695-1705.

Lee, E., C. K. Chan, and P. Paatero (1999). "Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong." *Atmospheric Envronment* 33, 3201– 3212.

Leemans, R., and Cramer, W.P. (1992). "IIASA Database for Mean Monthly Values of Temperature, Precipitation, and Cloudiness on a Global Terrestrial Grid. Digital Raster Data on a 30 minute Cartesian Orthonormal Geodetic (lat/long) 360x720 grid." In: *Global Ecosystems Database Version 2.0*. Boulder, CO: NOAA National Geophysical Data Center. Thirty-six independent single-attribute spatial layers. 15,588,254 bytes in 77 files. [first published in 1991]

Lee-Taylor, J. M., G. P. Brasseur, and Y. Yokouchi (2001). "A preliminary threedimensional global model study of atmospheric methyl chloride distributions." *Journal of Geophysical Research* 106: 34,221– 34,233.

Levy, H., II, J. D. Mahlman, W. J. Moxim, and S. C. Liu (1985). "Tropospheric ozone: The role of transport." *Journal of Geophysical Research* 90, 3753– 3772.

Li, Q. B. *et al*. (2000). "Atmospheric Hydrogen Cyanide (HCN): Biomass burning source, Ocean sink?" *Geophysical Research Letters 27,* 357-360.

Li, Q. B. *et al*. (2003). "A global three-dimensional model analysis of the atmospheric budget of HCN and CH_3CN: Constraints from aircraft and ground measurements." *Journal of Geophysical Research 108* (D21), 8827, doi:10.1029/2002JD003075.

Li, Q., D. J. Jacob, T. D. Fairlie, H. Liu, R. V. Martin, and R. M. Yantosca (2002), "Stratospheric versus pollution influences on ozone at Bermuda: Reconciling past analyses." *Journal of Geophysical Research* 107: (D22), 4611, doi:10.1029/ 2002JD002138.

Limbeck, A., M. Kulmala and H. Puxbaum (2003). "Secondary organic aerosol formation in the atmosphere via heterogeneous reaction of gaseous isoprene on acidic particles." *Geophysical Research Letters* 30: (19), 1996, doi:10.1029/2003GL017738.

Liu, H., D. J. Jacob, I. Bey, and R. M. Yantosca (2001). "Constraints from ^{210}Pb and ^{7}Be on wet deposition and transport in a global three-dimensional chemical tracer model driven by assimilated meteorological fields." *Journal of Geophysical Research* 106: 12,109–12,128.

Liu, H., D.J. Jacob, L.Y. Chan, S.J. Oltmans, I. Bey, R.M. Yantosca, J.M. Harris, B.N. Duncan, and R.V. Martin (2002*).* "Sources of tropospheric ozone along the Asian Pacific Rim: An analysis of ozonesonde observations." *Journal of Geophysical Research*, 10.1029/2001JD002005.

Liu, H., D.J. Jacob, J.E. Dibb, A.M. Fiore, and R.M. Yantosca (2004*).* "Constraints on the sources of tropospheric ozone from ^{210}Pb-^{7}Be-O_3 correlations.*"* *Journal of Geophysical Research* 109, D07306, doi:10.1029/2003JD003988.

Liu, S. C., M. Trainer, F. C. Fehsenfeld, D. D. Parrish, E. J. Williams, D. W. Fahey, G. Hu¨bler, and P. C. Murphy (1987). "Ozone production in the rural troposphere and the implications for regional and global ozone distributions." *Journal of Geophysical Research* 92: 4191– 4207.

Liu, W. (2002). "Application of receptor modeling to atmospheric consitituents at Potsdam and Stockton, NY." Clarkson University. Ph.D. Dissertation.

Liu, W., Y. Wang, A. Russell, and E. Edgerton (2005). "Atmospheric aerosols over two urban-rural pairs in southeastern United States: Chemical composition and sources." *Atmospheric Environment 39*: 4453-4470.

Liu, W. *et al*. (2005). "Atmospheric aerosol over two urban-rural pairs in the southestern United States: Chemical composition and sources." *Atmospheric Environment., 39,* 4,453**–**4,710.

Liu, W., Y. Wang, A. Murmur, A. Russell, and E. Edgerton (2006). "Evaluation of CMAQ simulated atmospheric constituents with SEARCH observations in the factor projected space.", submitted to *Journal of Geophysical Research.*

Logan, J. A. (1985). "Tropospheric ozone: Seasonal behavior, trends, and anthropogenic influence." *Journal of Geophysical Research* 90, 10,463– 10,482.

Lurmann, F. W., A. C. Lloyd, and R. Atkinson (1986). "A chemical mechanism for use long-range transport acid deposition computer modeling." *Journal of Geophysical Research* 91, 905–936.

Marandino, C. A., *et al*. (2005). "Oceanic uptake and the global atmospheric acetone budget." *Geophysical Research Letters 32,* L15806, doi:10.1029/2005GL023285.

Martin, R.V., K. Chance, D.J. Jacob, T.P. Kurosu, R.J.D. Spurr, E. Bucsela, J.F. Gleason, P.I. Palmer, I. Bey, A.M. Fiore, Q. Li, R.M. Yantosca, and R.B.A. Koelemeijer, (2002). "An improved retrieval of tropospheric nitrogen dioxide from GOME." *Journal of Geophysical Research* 107: (D20), 4437, doi:10.1029/2001JD001027.

Martin, R.V., D.J. Jacob, R.M. Yantosca, M. Chin, and P. Ginoux (2003). "Global and Regional Decreases in Tropospheric Oxidants from Photochemical Effects of Aerosols." *Journal of Geophysical Research* 108(D3), 4097, doi:10.1029/2002JD002622.

Mauzerall, D. L., D. Narita, H. Akimoto, L. Horowitz, S. Walters, D. A. Hauglustaine, and G. Brasseur (2000). "Seasonal characteristics of tropospheric ozone production and mixing ratios over East Asia: A global three dimensional chemical transport model analysis." *Journal of Geophysical Research* 105, 17,895– 17,910.

McCulloch, A. *et al*. (1999). "Global emissions of hydrogen chloride and chloromethane from coal combustion, incineration, and industrial activities: Reactive chlorine emissions inventory." *Journal of Geophysical Research 104*, 8,391–8,403.

Moxim, W.J., H. Levy II, and P.S. Kasibhatlan (1996). "Simulated global troposhperic PAN: Its transport and impact on NO_X." *Journal of Geosphysical Research* 101: 12,621-12,638.

Mueller, J.-F. (1992). "Geographical distribution and seasonal variation of surface emissions and deposition velocities of atmospheric trace gases" *Journal of Geosphysical Reserach*, 97: 3787-3804.

Olson, J., World ecosystems (WE1.4) (1992). "Digital raster data on a 10 minute geosgraphic 1080 x 2160 grid" in *Global ecosystems database, Version 1.0: Disc A*, edited by NOAA National Geophysical Data Center, Boulder, CO.

Orlando, J.J., B. Nozière, G. S. Tyndall, G. E. Orzechowska, S. E. Paulson, and Y. Rudich (2000). "Product studies of the OH-and ozone initiated oxidation of some monoterpenes." *Journal of Geophysical Research* 105: 11,561-11,572.

Oltmans, S. J., and H. Levy II (1992). "Seasonal cycle of surface ozone over the western North Atlantic." *Nature*, 358, 392–394.

Paatero, P. and U. Tapper (1993). "Analysis of different modes of factor analysis as least squares fit problems." *Chemometrics and Intelligent Laboratory Stystems* 18: 183 – 194.

Paatero, P. and U. Tapper (1994). "Positive Matrix Factorization: a non-negative factor model with optimal utilization of error estimates of data values." *Environmetrics* 5: 111 -126.

Paatero, P. (1997). "Least squares formulation of robust non-negative factor analysis." *Chemometrics Intelligent Laboratory System* 37, 23–35.

Paatero, P. (2000). "User's Guide for Positive Matrix Factorization programs PMF2 and PM3, part 2: reference". University of Helsinki, Helsinki.

Paatero, P., P. K. Hopke, X.-H. Song, and Z. Ramadan (2002). "Understanding and controlling rotations in factor analytic models." *Chemometrics Intelligent Laboratory System* 60: 253– 264.

Palmer, P. I., D. J. Jacob, K. Chance, R. V. Martin, R. J. D, Spurr, T. P. Kurosu, I. Bey, R. Yantosca, A. Fiore, and Q.B. Li (2001). "Air mass factor formulation for spectroscopic measurements from satellites: application to formaldehyde retrievals from GOME." *Journal of Geophysical Resarch* 106: 14,539-14,550.

Palmer, P. I., D. J. Jacob, A. M. Fiore, R. V. Martin, K. Chance, and T. P. Kurosu (2003). "Mapping isoprene emissions over North America using formaldehyde column observations from space." *Journal of Geosphysical Research* 108: doi:10.1029/2000JD002153.

Palmer, P. I., D. J. Jacob, D. B. A. Jones, C. L. Heald, R. M. Yantosca, J. A. Logan, G. W. Sachse, and D. G. Streets (2003). "Inverting for emissions of carbon monoxide from Asia using aircraft observations over the western Pacific." *Journal of Geophysical Research* 108: 8825, doi:10.1029/2002JD003176.

Paterson, K. G., J. L. Sagady, and D. L. Hooper (1999). "Analysis of air quality data using positive matrix factorization." *Environmental Science and Technology* 33, 635– 641.

Platt, U. (1994). "Differential Optical Absorption Spectrosopy (DOAS), in Air Monitoring by Spectroscopic Techniques, pp. 27 – 84, *John Wiley,* New York.

Pouton, J. E. (1990). "Cyanogenesis in plants." *Plant Physiology 94*, 401–405.

Penkett, S. A., and K. A. Brice (1986) "The spring maximum in photo-oxidants in the Northern Hemisphere troposphere." *Nature* 319: 655– 657.

Pierce, T., C. Geron, L. Bender, R. Dennis, G. Tonnesen, and A. Guenther (1998). "Influence of increased isoprene emissions on regional ozone modeling." *Journal of Geosphysical Research* 103: 25,611-25,629.

Pochanart, P., H. Akimoto, S. Maksyutov, and J. Staehelin (2001). "Surface ozone at the Swiss Alpine site Arosa: The hemispheric background and the influence of large-scale anthropogenic emissions." *Atmospheric Environment* 35, 5553– 5566.

Poisson, N., M. Kanankidou, and P.J. Crutzen (2000). "Impact of non-methane hydrocarbons on tropospheric chemistry and the oxidizing power of the global troposphere, 3-dimensional modeling results." *Journal of Atmospheric Chemstry* 36: 157-230.

Polissar, A. V., P. K. Hopke, P. Paatero, W. C. Malm, and J. F. Sisler (1998). "Atmospheric aerosol over Alaska: 2. Elemental composition and sources." *Journal of Geophysical Research* 103: 19,045– 19,057.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. F. Flannery (1992). "Numerical Recipes in Fortran" 963 pp., Cambridge Univ. Press, New York.

Prinn, R., *et al* (2001). "Evidence for substantial variations of atmospheric hydorxly radicals in the past two decades" *Science* 292: 1882-1888.

Ramadan, Z., X.-H. Song, and P. K. Hopke (2000). "Identification of source of Phoenix aerosol by positive matrix factorization." *Journal of Air Waste Management Association* 50, 1308– 1320.

Raper, J. L. *et al*. (2001). "Pacific Exploratory Mission in the tropical Pacific: PEM-Tropics B, March–April 1999." *Journal of Geophysical Research 106*, 32,401–32,425.

Rinsland, C. P. *et al*. (1998). "ATMOS/ATLAS 3 infrared profile measurements of trace gases in the November 1994 tropical and subtropical upper troposphere." *Journal of Quantatative Spectroscopy & Radiative Transfer, 60*, 891–901.

Rodgers, C.D. (2000). "Inverse Methods for Atmospheric Sounding: Theory and Practic." *World Sci.*, River Edge, N. J.

Schubert, S.D., R.B. Rood, and J. Pfaendtner (1993). "An assimilated data set for earth science applications." *Buletin of American Meteorological Society* 105: 19,991-20,011.

Singh, H. *et al*. (2000). "Distribution and fate of selected oxygenated organic species in the troposphere and lower stratosphere over the Atlantic" *Journal of Geophysical Research* 105: 3,795–3,805.

Singh, H., Chen, Y., Staudt, A., Jacob, D., Blake, D., Heikes, B., and Snow, J. (2001). "Evidence from the Pacific troposphere for large global sources of oxygenated organic compounds." *Nature* 410: 1078-1081.

Singh, H, D. Jacob, L. Pfister, and J. Crawford (2002). "INTEX-NA: Intercontinental Chemical Transport Experiment - North America." from http://cloud1.arc.nasa.gov/docs/intex-na/white_paper.pdf (May/2006).

Singh, H. *et al*. (2003a). "In situ measurements of HCN and CH_3CN over the Pacific Ocean: Sources, sinks, and budgets." *Journal of Geophysical Research 108* (D20), 8795, doi: 10.1029/2002JD003006.

Singh, H. *et al*. (2003b). "Oxygenated volatile organic chemicals in the ocean: Inferences and implications based on atmosphere observations and air-sea exchange models." *Geophysical Research Letters 30*, 16, 1862, doi: 10. 1029/2003GL017933.

Singh, H. *et al*. (2004). "Analysis of the atmospheric distribution, sources, and sinks of oxygenated volatile organic chemicals based on measurements over the Pacific during TRACE-P." *Journal of Geophysical Research., 109*, D15S07, doi:10.1029/2003JD003883.

Sokolik, I. (2004). "Remote sensing of the atmosphere and oceans (Lecture)." Dept. of Earth and Atmospheric Sciences, http://irina.eas.gatech.edu/Teaching.htm (May/2006).

Spivakovsky, C. M., J. A. Logan, S. A. Montzka, Y. J. Balkanski, M. Foreman-Fowler, D. B. A. Jones, L. W. Horowitz, A. C. Fusco, C. A. M. Brenninkmeijer, M. J. Prather, S. C. Wofsy, and M. B. McElroy (2000). "Three-dimensional climatological distribution of tropospheric OH: Update and evaluation" *Journal of Geophysical Research* 105: 8931-8980.

Spurr, R.J.D. (2002). "Simultaneous derivation of intensities and weighting functions in a general pseudo-spherical discrete ordinate radiative transfer treatment." *Journal of Quantitative Spectroscopy & Radiative Transfer* 75: 129-175.

Stohl, A. (1998). "Computation, accuracy and applications of trajectories—A review and bibliography." *Atmospheric Environment* 32, 947– 966.

Tanimoto, H., O. Wild, S. Kato, H. Furutani, Y. Makide, Y. Komazaki, S. Hashimoto, S. Tanaka, and H. Akimoto (2002). "Seasonal cycles of ozone and oxidized nitrogen species in northeast Asia: 2. A model analysis of the roles of chemistry and transport." *Journal of Geophysical Research* 107(D23), 4706 doi:10.1029/ 2001JD001497.

van Meurs, A. P. H. (1997). "World Fiscal Systems for Gas—1997." vol. 1, van Meurs and Association., New York.

Vetter, J. (2000). "Pant cyanogenic glycosides." *Toxicon, 38*, 11–36.

Wang, Y., and D.J. Jacob (1998). "Anthropogenic forcing on tropospheric ozone and OH since preindustrial times." *Journal of GeophysicalResearch* 103: 31,123-31,135.

Wang, Y., D.J. Jacob, and J.A. Logan (1998a). "Global simulation of tropospheric O_3-NOx-hydrocarbon chemistry, 1. Model formulation." *Journal of Geophysical Research* 103: D9, 10,713-10,726.

Wang, Y., D.J. Jacob, and J.A. Logan (1998b). "Global simulation of tropospheric O_3-NO_X-hydrocarbon chemistry, 3. origin of troposheric ozone and effects of non-methane hydrocarbons." *Journal of Geosphysical Research* 103: 10,757-10,767.

Wang, Y., *et al*. (2001). "Factors Controlling Tropospheric O_3, OH, NO_x, and SO_2 over the Tropical Pacific during PEM-Tropics B." *Journal of Geophysical Research 106*, 32,733-32,748.

Wang, Y. *et al*. (2003a). "Springtime photochemistry at northern mid and high latitudes." *Journal of Geophysical Research* 108: (D4), 8358, doi:10.1029/2002JD002227.

Wang, Y. *et al*. (2003b). "Intercontinental transport of pollution manifested in the variability and seasonal trend of springtime O_3 at northern middle and high latitudes." *Journal of Geophysical Research 108* (D21), 46 83, doi: 10.1029/2003JD003592.

Wang, Y., and T. Zeng (2004). "On tracer correlations in the troposphere: The case of ethane and propane." *Journal of Geophysical Research 109*, D24306, doi:10.1029/2004JD005023.

Wang, Y. *et al* (2006). "Late-spring increase of trans-Pacific pollution transport in the upper troposphere." *Geophysical Research Letters* 33, L01811, doi:10.1029/2005GL024975.

Wild, O., X. Zhu, and M. J. Prather (2000). "Fast-J: Accurate simulations of in- and below-cloud photolysis in tropospheric chemistry models." *Journal of Atmospheric Chemistry* 37, 245–282.

Wild, O., and H. Akimoto (2001). "Intercontinental transport of ozone and its precursors in a three-dimensional global CTM." *Journal of Geophysical Research* 106: 27,729–27,744.

Xie, Y.-L., P. K. Hopke, P. Paatero, L. A. Barrie and S.-M. Li (1999). "Identification of Source Nature and Seasonal Variation of Arctic Aerosol by Positive Matrix Factorization." Journal of Atmospheric Sciences 56: 249 – 260.

Yienger, J. J., A. A. Klonecki, H. Levy II, W. J. Moxim, and G. R. Carmichael (1999). "An evaluation of chemistry's role in the winter-spring ozone maximum found in the northern midlatitude free troposphere." *Journal of Geophysical Research* 104, 3655 – 3668, 1999. Correction, *Journal of Geophysical Research* 104: 8329–8330.

Yokouchi, Y., *et al*. (2000). "A strong source of methyl chloride to the atmosphere from tropical coastal land." *Nature*, 403, 295– 298.

Yoshida, Y. *et al*. (2004). "A three-dimensional global model study of atmospheric methyl chloride budget and distributions." *Journal of Geophysical Research 109*, D24309, doi: 24310.21029/2004JD004951.

Zanis, P., P. S. Monks, E. Schuepbach, and S. A. Penkett (2000). "The role of in situ photochemistry in the control of ozone during spring at the Jungfraujoch (3580 m asl): Comparison of model results with measurements." *Journal of Atmospheric Chemistry* 37: 1– 27.

Zimmerman, P. (1979). "Testing of hydrocarbon emissions from vegetation, leaf litter and aquatic surfaces, and development of a method for compiling biogenic emission inventories." Rep. EPA-450-4-70-004, U.S. Environmental Protection Agency, Research Triangle Park, N.C.